A Mathematical Model to Study the Joint Effects of Genetics and Diet on Obesity

Fangyuan Hong[1], Victoria Kelley[2], Kevin Molina-Serrano[3], Demetrius Rhodes[4], Daniel Burkow[5], Marlio Paredes[5], and Karen Rios-Soto[*3,5]

[1]Mount Holyoke College
[2]James Madison University
[3]University of Puerto Rico Mayaguez
[4]University of South Carolina Beaufort
[5]Arizona State University

September 24, 2015

Abstract

Obesity has become one of the most pervasive epidemics facing North America today. According to the Centers for Disease Control and Prevention (CDC), more than one-third (34.9%) of adults in the United States are obese, and approximately 17% of children and adolescents aged 2 - 19 years are obese. Obesity is correlated with other serious health threats such as diabetes and cardiovascular diseases that increase an individual's mortality risk. Previous studies show that a particular SNP (*rs9939609*) in the fat mass and obesity associated *FTO* gene is associated with the expression of obesity. A poor choice of diet and nutrition may lead to obesity. In this study, we build a system of non-linear ordinary differential equations that considers both genetic and environmental effects on populations with three distinct genotypes (AA, Aa and aa). The autosomal dominant allele is A, therefore individuals who have the genotypes AA and Aa express the *FTO* gene. Equilibria analysis and simulation results show that over a long period of time, when the birth frequency of each genotype is dependent on current allele frequencies, the proportion of populations with the dominant allele goes to 0, or the dominant allele A is outbred by the recessive gene allele. Simulation results further show that having the allele A has a stronger impact on obesity than the diet environment. Thus the effect of environmental factors on the dynamics of obesity are negligible at best. Fitness and genetic selection trumps any environmental bias. This study provides a significantly new insight into the synergic impact that genetics and diet play on obesity, which is rarely studied by traditional biological tools, such as

*karen.rios3@upr.edu

GWAS. Note that with genetic inheritance, environment makes no significant impact on the prevalence of obesity in the long term.

1 Introduction

Obesity is a complex disease involving an elevated accumulation of body fat which can increase the risk of many health problems. According to the Centers for Disease Control and Prevention (CDC), an individual having a body mass index (BMI) greater than or equal to 30 is classified as obese. Currently, 64% of the population of the United States is estimated to be overweight or obese [9]. It is prevalent among individuals of both genders, all levels of socio-economic status, and all ethnic groups [11]. No country has recorded any attenuation or reversal of the epidemic [14]. In the United States, obesity prevalence in 2013 varied across states and territories; the South had the highest prevalence of obesity (30.2%), followed by the Midwest (30.1%), the Northeast (26.5%) and the West (24.9%) [5].

Obesity has several major comorbid health effects, including diabetes, cardiovascular heart diseases, various cancers, and arthritis, which are expensive to treat and may decrease an individual's life expectancy [14]. Infectious diseases and nutrient deficiency diseases are thus being replaced by new threats such as obesity, diabetes and cardiovascular diseases [14]. In fact, diabetes is rapidly emerging as a global health threat that may reach a pandemic level by 2030 [8]. Hence it is clear that obesity could lead to a high risk of all-cause mortality [1].

Emerging studies indicate that genetic factors may alter the magnitude of weight loss and lipid change in response to behavioral treatment [9]. The effect of genetics on the pathogenesis of obesity can be intuitively explained as an intrinsic resistance to the positive changes typically generated by a healthy lifestyle [9]. Genetic factors play a critical role in explaining how some patients' metabolisms prevent them from responding to lifestyle changes in an environment that is effective for other people [9]. Therefore, genetics is an important factor to consider when studying an obesity model [9, 10, 12].

In addition to genetics, poor diet and nutrition choices are strongly associated with obesity [14]. The increasing urbanization and westernization occurring in most countries around the world is associated with changes in the diet towards one of high fat, high energy-density foods and a sedentary lifestyle [14]. This shift is also associated with the current rapid changes in childhood and adult obesity [4, 14]. Previous studies have shown that the main food-related vector that promotes passive over-consumption of calories is a high intake of energy dense food such as many processed foods [14]. On the other hand, a healthy diet to fight and avoid obesity is one with high dietary non-starch polysacchrides (NSP) or fiber intake.

The *FTO* (fat mass and obesity associated) gene was first discovered in a genome-wide association study (GWAS) in 2007 [11]. Studies have shown a significant association between five *FTO* polymorphisms and obesity risk across different ethnic groups [11]. Among them, the single-nucleotide polymorphism (SNP) *rs9939609* has been of particular interest as it displays a relatively consistent association with obesity across multiple ethnic groups [11]. However, the mechanisms that trigger obesity by the *FTO* gene are unclear. In this work, we will only consider this single SNP (*rs9939609*) as a potential trigger of obesity. We include this in our model by using a single pair of alleles.

Multiple studies have confirmed the positive correlation between having the particular SNP in the *FTO* gene. Longitudinal measurements were assessed in one study of twenty-six individuals, from 1976-2002 [2]. A follow up study was conducted on the cohorts – 26 years for the women and 16 years for the men – to confirm if there was an association between having the particular SNP, rs9939609, and adiposity [2]. It was found that the dominant AA genotype (those with the SNP) was associated with greater obesity risk during the follow-up in women and men [2]. Another study confirmed that high-fat diets and low physical activity levels may accentuate the susceptibility to obesity in those with the FTO variant [12]. After each of the cohorts were genotyped and their BMI was collected [12], it was discovered that the individuals with the dominant AA had a higher BMI than the individuals who had the recessive aa genotype [12]. Since a high intake of fat was correlated with a low level of leisure-time physical activity in this study, the BMI differences across FTO genotypes with different fat intakes and physical activity levels were examined [12].

Ejima et al. (2015) used a susceptible-infected-recovered (SIR) model framework to study the social contagion dynamics of obesity [7]. The epidemiological process of becoming and recovering from obesity were functions of time. To capture the population dynamics, ordinary differential equations were used to describe the time dependence obesity risk [7]. On the other hand, Thomas et al. (2014) used an SIR model with six ordinary differential equations describing the interaction and transitions between populations [15]. The individuals were divided into infected and non-infected classes where individuals with a BMI below 25 were classified as susceptible [15]. Both works studied obesity solely as a social dynamics process without taking into consideration genetic factors.

In their book, Brauer and Kribs (2015) examined a model with slow selection in population genetics under the Hardy-Weinberg principle [3]. Assuming equally fit genotypes, density-dependent population dynamics and a genotype-dependent death rate, the book chapter developed genotype frequencies, and proportions of homozygous dominant, heterozygous dominant and homozygous recessive in terms of state variables, expressions that we will adapt for our model. However, in our proposed model, selection happens as a result of obese and non-obese individuals in the whole population mixing homogeneously and randomly.

In Section 2 we look at the complete ideal model. Section 2.1 outlines our assumptions for a simplified model that is analyzed and studied in this work. Section 3 is our first scenario where we consider a linear transmission of people moving into the obese class without recovery. This is divided into two cases depending on the allele frequencies. Section 4 is our second scenario with a nonlinear transfer rate that incorporates diet and environmental factors to and from the obese class. We also divide this section depending on the allele frequencies. We then compare the results of the model (Section 5), discuss the biological relevance in our conclusion (Section 6), and discuss future work (Section 7).

2 Full Model

In order to fill the gap between the aforementioned population and biological studies on the genetics of obesity and the interaction between genetic factors and environmental stimuli, we build a deterministic model that considers vertical transmission of alleles and differential fitness between obese and non-obese people to study changes in allele frequencies.
Many population genetic models assume a deterministic process to changes in allele frequencies [6]. Our deterministic model only considers vertical transmission of genes, and neglect the possible mutation or migration of genes, which may cause horizontal transfer of genotypes at one generation. The Hardy-Weinberg Principle of genetics states that the frequencies of genes and alleles tend to remain fairly constant from one generation to the next if individuals in a large population mix homogeneously and have equal fitness [3]. We predict that the proportion of obesity among the total population would change with time, and the gene frequency of obesity-inducing allele (A) would decrease and reach an equilibrium in the long run. Our model considers a population with density-dependent and genotype-independent birth and death rates respectively, whose dynamics shift in a continuous-time frame [3]. Based on the Hardy-Weinberg Principle of genetics, we define p to be the frequency of the dominant A allele (those with SNP rs9939609) and that q is the frequency of the recessive allele a (those without SNP rs9939609) ; thus $p + q = 1$.

We formulate a general model that incorporates both the genetics and environmental effects on obesity. This model consists of a system of 9 non-linear ordinary differential equations. The total population is divided into three categories: people who eat a healthy diet and are not obese (H), those who eat an unhealthy diet and are not yet obese (U), and people who are obese (O). The healthy diet is defined as a high fiber intake, low fat cholesterol sugar diet; while the unhealthy diet is referred to a high intake of fat protein energy-dense foods, and a low fiber intake. Each of these categories is divided into three sub categories based on individual's genotype: AA, Aa and aa. The dominant allele (A) represents the FTO gene variant with the rs9939609 SNP, meaning that the allele can be inherited if the parent genotype is either AA or Aa (Table 1 for further explanation of the state variables).

We make a variety of assumptions to construct our general model. Individuals are either born into a healthy or unhealthy environment at the rate α, with a proportion r of healthy individuals in the population and a proportion $(1 - r)$ of unhealthy individuals in the population. The transfer function ξ_i [1] from healthy to unhealthy eaters, the obesity rate β_{ij} [2] from healthy eaters to obese people, and the transfer function ρ_i from unhealthy to healthy eaters are different for people with genotypes AA and Aa than for those with genotype aa. That is, we assume people who have the SNP variant within the FTO gene care more about their diet than those without the gene. To recover from obesity, one has to adopt a heathy diet, and people with or without the FTO gene variant have different innate resistance to recover. Hence the recovery rate γ_i from obese people to healthy eaters is also different for people with the genotype AA and Aa than for those with genotype aa. Similarly, the progression rate ϕ_i from unhealthy eaters to obesity is different. In addition, people have different death rates μ_j [3] depending on their genotype and diet. Since obesity may lead to a higher mortality rate, death rates for those in the obese class are greater than those in healthy and unhealthy classes.

[1] $i \in \{1, 2\}$, where 1 refers to genotype AA, and 2 refers to genotype aa.

[2] $i \in \{1, 2\}$ referring to genotype AA/Aa and aa respectively, and $j \in \{1, 2\}$ referring to linear and nonlinear function respectively.

[3] $j \in \{1, 2, 3, ..., 9\}$

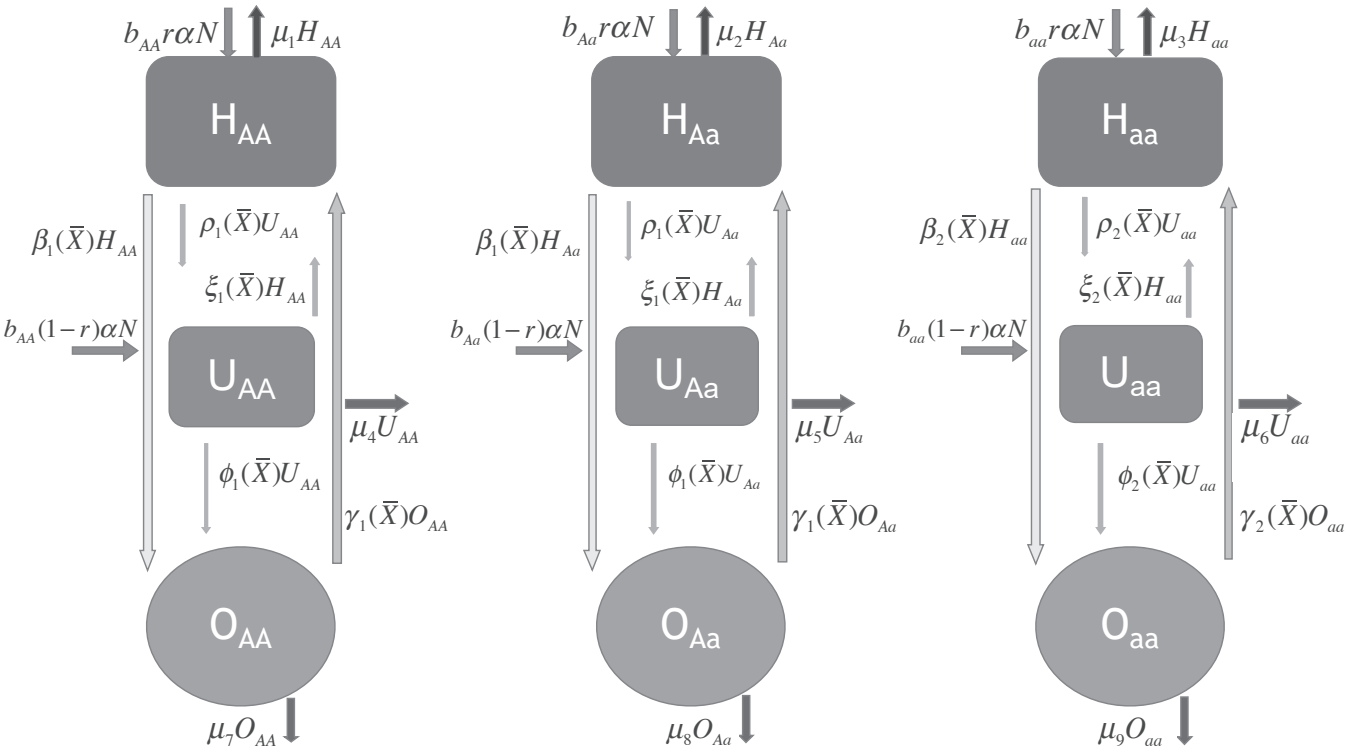

Figure 1: Flow chart of full model.

If we let,

$$\bar{X} \quad = \quad [H_{AA}, H_{Aa}, H_{aa}, U_{AA}, U_{Aa}, U_{aa}, O_{AA}, O_{Aa}, O_{aa}],$$

then the full model (see Figure 1) is given by the following system of non-linear differential

equations:

$$
\begin{aligned}
H'_{AA} &= b_{AA} r\alpha N - \xi_1\left(\bar{X}\right) H_{AA} + \rho_1\left(\bar{X}\right) U_{AA} - \beta_1\left(\bar{X}\right) H_{AA} + \gamma_1\left(\bar{X}\right) O_{AA} - \mu_1 H_{AA}, && \text{(1a)}\\
H'_{Aa} &= b_{Aa} r\alpha N - \xi_1\left(\bar{X}\right) H_{Aa} + \rho_1\left(\bar{X}\right) U_{Aa} - \beta_1\left(\bar{X}\right) H_{Aa} + \gamma_1\left(\bar{X}\right) O_{Aa} - \mu_2 H_{Aa}, && \text{(1b)}\\
H'_{aa} &= b_{aa} r\alpha N - \xi_2\left(\bar{X}\right) H_{aa} + \rho_2\left(\bar{X}\right) U_{aa} - \beta_2\left(\bar{X}\right) H_{aa} + \gamma_2\left(\bar{X}\right) O_{aa} - \mu_3 H_{aa}, && \text{(1c)}\\
U'_{AA} &= b_{AA}\left(1-r\right)\alpha N + \xi_1\left(\bar{X}\right) H_{AA} - \rho_1\left(\bar{X}\right) U_{AA} - \phi_1\left(\bar{X}\right) U_{AA} - \mu_4 U_{AA}, && \text{(1d)}\\
U'_{Aa} &= b_{Aa}\left(1-r\right)\alpha N + \xi_1\left(\bar{X}\right) H_{Aa} - \rho_1\left(\bar{X}\right) U_{Aa} - \phi_1\left(\bar{X}\right) U_{Aa} - \mu_5 U_{Aa}, && \text{(1e)}\\
U'_{aa} &= b_{aa}\left(1-r\right)\alpha N + \xi_2\left(\bar{X}\right) H_{aa} - \rho_2\left(\bar{X}\right) U_{aa} - \phi_2\left(\bar{X}\right) U_{aa} - \mu_6 U_{aa}, && \text{(1f)}\\
O'_{AA} &= \beta_1\left(\bar{X}\right) H_{AA} - \gamma_1\left(\bar{X}\right) O_{AA} + \phi_1\left(\bar{X}\right) U_{AA} - \mu_7 O_{AA}, && \text{(1g)}\\
O'_{Aa} &= \beta_1\left(\bar{X}\right) H_{Aa} - \gamma_1\left(\bar{X}\right) O_{Aa} + \phi_1\left(\bar{X}\right) U_{Aa} - \mu_8 O_{Aa}, && \text{(1h)}\\
O'_{aa} &= \beta_2\left(\bar{X}\right) H_{aa} - \gamma_2\left(\bar{X}\right) O_{aa} + \phi_2\left(\bar{X}\right) U_{aa} - \mu_9 O_{aa}, && \text{(1i)}
\end{aligned}
$$

where $\dfrac{dX}{dT} \equiv X'$, and $N = H_{AA} + H_{Aa} + H_{aa} + U_{AA} + U_{Aa} + U_{aa} + O_{AA} + O_{Aa} + O_{aa}$ is the total population. Table 1 provides the explanation of state variables while Table 2 provides a definition for each of the parameters in our model.

Table 1. Explanation of state variables

State Variable	Meaning
H_{AA}	population of healthy eaters with genotype AA
H_{Aa}	population of healthy eaters with genotype Aa
H_{aa}	population of healthy eaters with genotype aa
U_{AA}	population of unhealthy eaters with genotype AA
U_{Aa}	population of unhealthy eaters with genotype Aa
U_{aa}	population of unhealthy eaters with genotype aa
O_{AA}	population of obese people with genotype AA
O_{Aa}	population of obese people with genotype Aa
O_{aa}	population of obese people with genotype aa

Table 2. Explanation of parameter functions.

Parameter	Meaning
r	proportion of people who are born into a healthy environment
α	natural birth rate $\left(\text{years}^{-1}\right)$
$\beta_1(\bar{X})$	obesity function transfer from the healthy to obesity with the FTO gene $\left(\text{years}^{-1}\right)$
$\beta_2(\bar{X})$	obesity function transfer from the healthy to obesity without the FTO gene $\left(\text{years}^{-1}\right)$
$\phi_1(\bar{X})$	obesity function transfer from the unhealthy to obesity with the FTO gene $\left(\text{years}^{-1}\right)$
$\phi_2(\bar{X})$	obesity function transfer from the healthy to obesity without the FTO gene$\left(\text{years}^{-1}\right)$
$\rho_1(\bar{X})$	conversion function from the unhealthy to the healthy with the FTO gene$\left(\text{years}^{-1}\right)$
$\rho_2(\bar{X})$	conversion function from the healthy to the unhealthy without the FTO gene$\left(\text{years}^{-1}\right)$
$\gamma_1(\bar{X})$	recovery function from obesity to the healthy with genotypes AA andAa $\left(\text{years}^{-1}\right)$
$\gamma_2(\bar{X})$	recovery function from obesity to the healthy with genotypes aa $\left(\text{years}^{-1}\right)$
μ_j	per capita death rate $\left(\text{years}^{-1}\right)$; $j = 1, \ldots, 9$
$b_{AA}(\bar{X})$	proportion of individuals born with genotype AA
$b_{Aa}(\bar{X})$	proportion of individuals born with genotype Aa
$b_{aa}(\bar{X})$	proportion of individuals born with genotype aa
$\xi_1(\bar{X})$	transfer function from the healthy eaters to unhealthy eaters with the FTO gene variant (genotype AA or Aa) $\left(\text{years}^{-1}\right)$
$\xi_2(\bar{X})$	transfer function from the healthy eaters to unhealthy eaters without the FTO gene variant (genotype aa) $\left(\text{years}^{-1}\right)$

We define the genotype frequencies as $S_{AA} = \frac{H_{AA}+U_{AA}+O_{AA}}{N}$ for homozygous dominant genotype; $S_{Aa} = \frac{H_{Aa}+U_{Aa}+O_{Aa}}{N}$ for heterozygous genotype; and $S_{aa} = \frac{H_{aa}+U_{aa}+O_{aa}}{N}$ for homozygous recessive genotype, so that $S_{AA} + S_{Aa} + S_{aa} = 1$.

Inheritance law in genetics states that the dominant allele frequency is calculated from the frequency of dominant genotype plus the half frequency of heterozygous genotype [3]. Therefore, we define allele frequencies $p = \frac{2S_{AA}+S_{Aa}}{2} = S_{AA} + \frac{S_{Aa}}{2}$ and $q = \frac{S_{Aa}+2S_{aa}}{2} = S_{aa} + \frac{S_{Aa}}{2}$ (both non-constant), where p is the frequency of allele A, q is the frequency of allele a, $p + q = 1$, and $p^2 + 2pq + q^2 = 1$. Due to the complicated nature of the full model, we build a simplified model with only the structure we need to answer our research question.

2.1 Simplified Model

For simplification, we use the following assumptions on the full model (Model 1): individuals who have the FTO gene variant and consume an unhealthy diet immediately become obese; therefore $U_{AA} = U_{Aa} = 0$. The only population can become obese are those with the FTO gene variant (AA or Aa) and only if they eat unhealthy food; that is, $O_{aa} = 0$. We also assume people without the gene variant eat unhealthy because they are not concerned

with their diet; hence $H_{aa} = 0$. We assume the death rate of non-obese is μ_1 and the death rate for obese people is μ_2. Since obesity can cause a higher mortality rate, we further assume that $\mu_2 > \mu_1$.

We will examine four cases that consider environment and birth ratios (See Table 3):

Table 3. Explanation of Birth and Transition Functions.

No Environment	**Environment**
$\beta(\hat{X}) = \hat{\beta}$ $\gamma(\hat{X}) = 0$	$\beta(\hat{X}) = \hat{\beta}(O_{AA} + O_{Aa} + U_{aa})/N$ $\gamma(\hat{X}) = \hat{\gamma}(H_{AA} + H_{Aa})/N$
Constant Birth Ratio	**Non-Constant Birth Ratio**
$b_{AA} = \hat{p}^2$ $b_{Aa} = 2\hat{p}\hat{q}$ $b_{aa} = \hat{q}^2$	$p = (2H_{AA} + 2O_{AA} + H_{Aa} + O_{Aa})/2N$ $q = (2U_{aa} + H_{Aa} + O_{Aa})/2N$ $b_{AA} = p^2$ $b_{Aa} = 2pq$ $b_{AA} = q^2$

- Case 1: None Environmental Effect, Constant Birth Ratio
- Case 2: None Environmental Effect, Non-Constant Birth Ratio (State Variable Dependent)
- Case 3: Environmental Effect, Constant Birth Ratio
- Case 4: Environmental Effect, Non-Constant Birth Ratio (State Variable Dependent)

Using these assumptions, our simplified model becomes:

$$U'_{aa} = b_{aa}\alpha N - \mu_1 U_{aa}, \tag{2a}$$

$$H'_{AA} = b_{AA}\alpha N - \beta_1\left(\hat{X}\right) H_{AA} + \gamma_1\left(\hat{X}\right) O_{AA} - \mu_1 H_{AA}, \tag{2b}$$

$$H'_{AA} = b_{Aa}\alpha N - \beta_1\left(\hat{X}\right) H_{Aa} + \gamma_1\left(\hat{X}\right) O_{Aa} - \mu_1 H_{Aa}, \tag{2c}$$

$$O'_{AA} = \beta_1\left(\hat{X}\right) H_{AA} - \gamma_1\left(\hat{X}\right) O_{AA} - \mu_2 O_{AA}, \tag{2d}$$

$$O'_{Aa} = \beta_1\left(\hat{X}\right) H_{Aa} - \gamma_1\left(\hat{X}\right) O_{Aa} - \mu_2 O_{Aa}, \tag{2e}$$

where $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$.

2.2 Parameter Values

The parameters in this model include α, μ_1, $\hat{\beta}$, and μ_2 (Table 3).

Values for allele frequencies, $\hat{p}$ and $\hat{q}$ were taken from the Genome Wide Associated Study (GWAS) [13]. We obtain numerical values for α and μ_1 from the CDC [5]. We calculate μ_2 using the CDC data that shows in extreme cases, people who are obese die up to fourteen years earlier than people who are not. We use simulations, knowing the obese population was 30% in 2015 to calculate our $\hat{\beta}$ and $\hat{\gamma}$.
Initial values are estimated based on data provided by a study on the *FTO* gene in cohorts of Danish men [16]. We assume the total population to start with is 100000. The proportion of obesity individuals among the total population of each genotype gives O_{AA}=12410 and O_{Aa}=21990, respectively. Proportions of H_{AA}=9030 and H_{Aa}=27210 can be estimated using $N_{AA}-O_{AA}$ and $N_{Aa}-O_{Aa}$ respectively. The rest of the population goes to U_{aa}=29360.

Table 3: Calculated parameters for genetics only model with constant b_{AA}, b_{Aa}, b_{aa}, and $\beta_1\left(\hat{X}\right)$

Parameter	Meaning	Calculated Value
$\hat{p}$	allele frequency of A	0.46
$\hat{q}$	allele frequency of a	0.54
α	per capita birth rate (years^{-1})	0.0124
μ_1	per capita death rate (years^{-1}) of non-obese people	0.0082
μ_2	per capita death rate (years^{-1}) of obese people	0.0156
$\hat{\beta}$	linear transfer function from healthy eaters to obesity with genotype AA or Aa	Varies
$\hat{\gamma}$	linear transfer function from obesity to healthy eaters with genotype AA or Aa	Varies

3 Scenario 1: Only Gene Factor

We simplified Model 2 to consider only the genetics factor on obesity (Figure 2). That is, the transfer function β becomes $\beta_1 = \hat{\beta}$. Since under this scenario genetics makes people obese, obese people cannot recover from obesity; thus $\gamma_1 = 0$. Here, the transfer rate β_1 from eating a healthy diet to becoming obese is the same for both AA and Aa genotype.

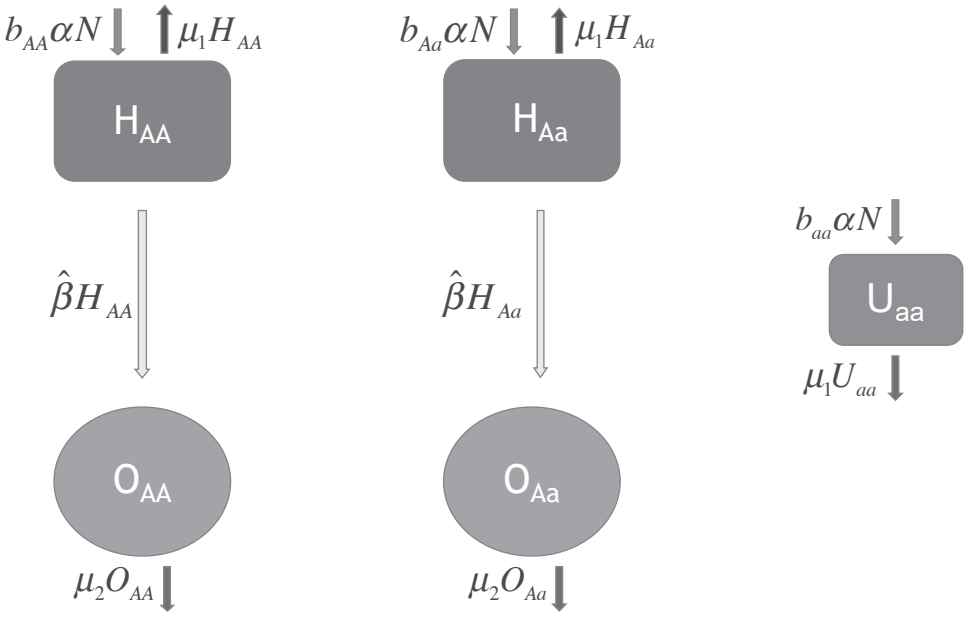

Figure 2: Flow chart of the reduced model with gene factor only.

3.1 Case 1: Analysis for Constant Birth Ratios

We first consider $b_{AA} = \hat{p}^2$, $b_{Aa} = 2\hat{p}\hat{q}$, and $b_{aa} = \hat{q}^2$; that is, the birth proportion of each genotype is constant. Recall that Model 2 reduces to

$$\begin{aligned} U'_{aa} &= b_{aa}\alpha N - \mu_1 U_{aa}, && \text{(3a)} \\ H'_{AA} &= b_{AA}\alpha N - \hat{\beta} H_{AA} - \mu_1 H_{AA}, && \text{(3b)} \\ H'_{Aa} &= b_{Aa}\alpha N - \hat{\beta} H_{Aa} - \mu_1 H_{Aa}, && \text{(3c)} \\ O'_{AA} &= \hat{\beta} H_{AA} - \mu_2 O_{AA}, && \text{(3d)} \\ O'_{Aa} &= \hat{\beta} H_{Aa} - \mu_2 O_{Aa}. && \text{(3e)} \end{aligned}$$

Note that $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$ is not constant, since $N' = (\alpha - \mu_1) N - (\mu_2 - \mu_1)(O_{AA} + O_{Aa})$. Model 3 is clearly linear, so we can analyze the single equilibrium. We proceed to analyze Model 3 by setting the birth ratios to be constant, thus $b_{AA} = \hat{p}^2$, $b_{Aa} = 2\hat{p}\hat{q}$, $b_{aa} = \hat{q}^2$. This means that the allele frequency does not update with the current population. Under this strong assumption, Model 3 becomes:

$$
\begin{aligned}
U'_{aa} &= \hat{q}^2 \alpha N - \mu_1 U_{aa}, && \text{(4a)}\\
H'_{AA} &= \hat{p}^2 \alpha N - \hat{\beta} H_{AA} - \mu_1 H_{AA}, && \text{(4b)}\\
H'_{Aa} &= 2\hat{p}\hat{q}\alpha N - \hat{\beta} H_{Aa} - \mu_1 H_{Aa}, && \text{(4c)}\\
O'_{AA} &= \hat{\beta} H_{AA} - \mu_2 O_{AA}, && \text{(4d)}\\
O'_{Aa} &= \hat{\beta} H_{Aa} - \mu_2 O_{Aa}, && \text{(4e)}
\end{aligned}
$$

where $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$.
Here, the change in population is given by the equation $N' = (\alpha - \mu_1) N - (\mu_2 - \mu_1) O$, where O is the total obese population. Notice that in the case we have $\mu_1 = \mu_2$, $N' = (\alpha - \mu_1) N$, that is $N(t) = N_0 e^{(\alpha - \mu_1)t}$ where N_0 is the total population at time zero.

Before proceeding to find equilibrium and determine stability of the Model 3, we first show positive invariance that proves that all state variables stay in the first quadrant.

3.1.1 Positive Invariance

We begin by showing that the Case 1 system is positively invariant. That is, if the initial conditions for all of the state variables are contained in the positive quadrant, then the forward time dynamics of every variable remains in the positive quadrant for all $t > 0$. We show this by contradiction.

Assume that $t_0 = 0$, $\hat{X}(0) \equiv \hat{X}_0 \geq 0$, $\alpha, \hat{\beta}, \hat{p}, \hat{q} > 0$, and that one of the variables is the first to become negative at some point on its trajectory. In order for this to occur, there must be:
i) a first time, t_c, that it leaves the positive quadrant,
ii) it must cross 0, and
iii) it must cross 0 with a negative derivative.

Suppose that $U_{aa}(t_c) = 0$, then

$$U'_{aa}(t_c) = U'_{aa}|_{U_{aa}=0} = \hat{q}^2 \alpha N.$$

Since U_{aa} is the first variable to cross outside of the positive quadrant we know $N \geq 0$. Thus we have $U'_{aa}(t_c) \geq 0$ which contradicts (iii).
Further, the same analysis can be applied to show that H_{AA} and H_{Aa} cannot be the first variable to cross outside of the positive quadrant. Then by extension, if we assume that O_{AA} is the first variable to leave the positive quadrant, $O_{AA}(t_c) = 0$ and we have

$$O'_{AA}(t_c) = O'_{AA}|_{O_{AA}=0} = \beta H_{AA}.$$

By our previous assertion that $H_{AA} \geq 0$ for all time, $t > 0$, we know that $O'_{AA}(t_c) \geq 0$ which contradicts (iii). The same analysis holds true for O_{Aa} given that $H_{Aa} \geq 0$. None of the five variables are the first to leave the positive quadrant, therefore none of the variables ever leave the positive quadrant for all $t > 0$, and our system is positively invariant.

3.1.2 Equilibrium

The only equilibrium for this linear model is $(0, 0, 0, 0, 0)$. Biologically, this means the population can either go extinct or grow exponentially when p and q are constant. The population will go extinct (the point will be stable) if the birth rate α is less than the death rate μ_1.

Theorem 1 *If $\alpha < \mu_1$ and $\mu_2 > \mu_1$, then our zero equilibrium in System 4 is globally asymptotically stable.*

Prove: Stability of the zero-equilibrium. The Jacobian matrix for Model 4 at the zero equilibrium is given by:

$$\mathbf{J} = \begin{pmatrix} q^2\alpha - \mu_1 & q^2\alpha & q^2\alpha & q^2\alpha & q^2\alpha \\ p^2\alpha & p^2\alpha - \beta - \mu_1 & p^2\alpha & p^2\alpha & p^2\alpha \\ 2pq\alpha & 2pq\alpha & 2pq\alpha - \beta - \mu_1 & 2pq\alpha & 2pq\alpha \\ 0 & \beta & 0 & -\mu_2 & 0 \\ 0 & 0 & \beta & 0 & -\mu_2 \end{pmatrix},$$

where its characteristic equation is provided by taking its determinant, that is

$$\begin{aligned} \det(\mathbf{J} - \lambda\mathbf{I}) &= (\beta + \mu_1 + \lambda)(\mu_2 + \lambda) \\ &\quad \Big[(\mu_1 + \lambda)^2(\alpha - \beta - \mu_1 - \lambda) + (\mu_1 + \lambda)((\mu_2 - \mu_1)(\alpha - \beta - \mu_1 - \lambda) + \beta\alpha) + \beta q^2 \alpha(\mu_2 - \mu_1)\Big] = 0. \quad (5) \end{aligned}$$

Two of the eigenvalues of Equation 5, $\lambda_1 = -\beta - \mu_1$ and $\lambda_2 = -\mu_2$ are clearly negative. In order to analyze the stability of the zero equilibrium further we need to know whether the other three eigenvalues are negative. In that case we have to look for the solutions of the cubic equation

$$\begin{aligned} 0 &= \lambda^3(-1) + \lambda^2(-3\mu_1 + \alpha - \beta - (\mu_2 - \mu_1)) \\ &\quad + \lambda\left(-3\mu_1^2 + 2\mu_1(\alpha - \beta - (\mu_2 - \mu_1)) + (\mu_2 - \mu_1)(\alpha - \beta) + \beta\alpha\right) \\ &\quad + \left[-\mu_1^3 + \mu_1^2(\alpha - \beta - (\mu_2 - \mu_1)) + \mu_1((\mu_2 - \mu_1)(\alpha - \beta) + \beta\alpha - 2) + \beta q^2 \alpha(\mu_2 - \mu_1)\right] \\ &= \lambda^3(-1) + \lambda^2(a_1) + \lambda(a_2) + a_3. \quad (6) \end{aligned}$$

Calculations for this are found in the Appendix under Section **??**. Setting $a_1, a_2, a_3 < 0$,

$$
\begin{aligned}
\alpha &< \mu_1 + \mu_1 + \mu_2 + \beta, \\
\alpha &< \mu_1 + \left(\frac{\mu_1 (\mu_2 + \beta) + (\mu_2 - \mu_1) \beta}{(\mu_1 + \mu_2 + \beta)} \right), \\
\alpha &< \mu_1 + \left(\frac{(\mu_2 - \mu_1) \beta \left(p^2 + 2pq\right) + 2}{\mu_1^2 + \mu_1 ((\mu_2 - \mu_1) + \beta) + \beta q^2 (\mu_2 - \mu_1)} \right) \mu_1.
\end{aligned}
$$

By Descartes' Rule of Signs, the equation does not have positive roots, and we know $\lambda \neq 0$ since there is a constant term in our equation. That is, all of the coefficients of Equation 6 are negative, provided $\alpha < \mu_1$ and $\mu_1 < \mu_2$. Thus the three remaining eigenvalues are negative. Therefore, our zero equilibrium is stable if $\alpha < \mu_1$ and $\mu_1 < \mu_2$. Biologically, this means our populations will approach zero, or go extinct, if the natural birth rate is less than the natural death rate.

3.1.3 Simulations

In this case, we calculated through simulations a $\hat{\beta}$ value that makes the obesity population reaches equilibrium at approximately 30% of the total population, the proportion suggested by the CDC data [5]. Using the values provided in Table 3 with $\hat{\beta}$ estimated to be 0.015, the simulation plots for Case 1 are in Figure 3. Notice that here $\alpha > \mu_1$, thus the extinction equilibrium is unstable.

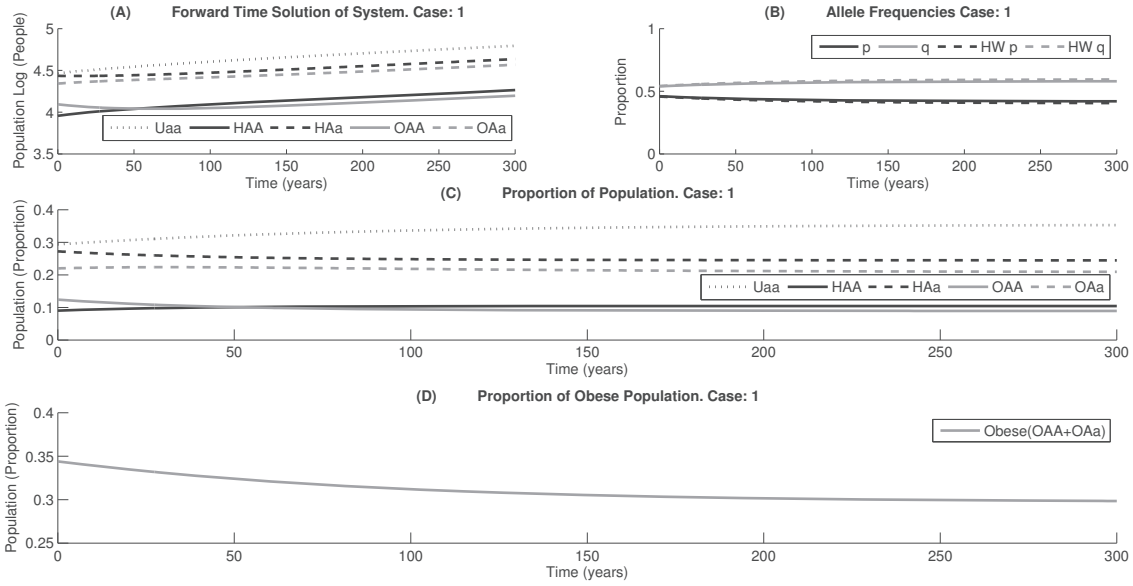

Figure 3: Simulation plots for Case 1.

Plot (A) on the upper left corner shows linear changes of log population with time (years). The population increases exponentially (when $\alpha > \mu_1$), so we use the logarithmic function to see the dynamics of the populations. The unhealthy population with genotype U_{aa} starts off at a large initial size because it has the largest initial population (Table 3). Plot (B) in the upper right corner shows changes of allele frequencies p and q with time. The real p and q are compared to theoretical values predicted by Hardy-Weinberg Principle, and suggest that gene/alleles frequencies converge to an equilibrium quickly. Our estimates of p and q match the theoretical values well. Within a short period of time, the recessive allele a (one without the FTO gene) has a greater proportion than the dominant allele A. Plot (C) shows changes of proportions of populations over time. Plot (D) shows changes of total obese population over time. The obese population decreases gradually and reaches constant level in about 200 years to about 30% of the total population.

3.2 Case 2: Analysis for Non-Constant Birth Ratios

For this case, birth proportions b_{XX} are not constant; that is, they depend on the state variables. This is important because in real life, gene proportions change depending on the current population. These assumptions give rise to Model 4 below,

$$\begin{aligned}
U'_{aa} &= b_{aa}\alpha N - \mu_1 U_{AA}, && \text{(7a)}\\
H'_{AA} &= b_{AA}\alpha N - \beta_1 H_{AA} - \mu_1 H_{AA}, && \text{(7b)}\\
H'_{Aa} &= b_{Aa}\alpha N - \beta_1 H_{Aa} - \mu_1 H_{Aa}, && \text{(7c)}\\
O'_{AA} &= \beta_1 H_{AA} - \mu_2 O_{AA}, && \text{(7d)}\\
O'_{Aa} &= \beta_1 H_{Aa} - \mu_2 O_{Aa}, && \text{(7e)}
\end{aligned}$$

where $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$.
Since the genotype frequencies are not constant, they are now changed to $S_{AA} = \frac{H_{AA}+O_{AA}}{N}$, $S_{Aa} = \frac{H_{Aa}+O_{Aa}}{N}$, and $S_{aa} = \frac{U_{aa}}{N}$.

Therefore, our gene/allele frequencies now become:

$$
\begin{aligned}
b_{AA} &= p^2 = \left(S_{AA} + \frac{S_{Aa}}{2}\right)^2 \\
&= \left(\frac{H_{AA} + O_{AA}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)^2, && \text{(8a)} \\
b_{Aa} &= 2pq = 2\left(S_{AA} + \frac{S_{Aa}}{2}\right)\left(S_{aa} + \frac{S_{Aa}}{2}\right) \\
&= 2\left(\frac{H_{AA} + O_{AA}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)\left(\frac{U_{aa}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right), && \text{(8b)} \\
b_{aa} &= q^2 = \left(S_{aa} + \frac{S_{Aa}}{2}\right)^2 \\
&= \left(\frac{U_{aa}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)^2. && \text{(8c)}
\end{aligned}
$$

Plugging the Expressions 8 into Model 4, the genetics factor model becomes

$$
\begin{aligned}
U'_{aa} &= \left(\frac{U_{aa}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)^2 \alpha N - \mu_1 U_{aa}, && \text{(9a)} \\
H'_{AA} &= \left(\frac{H_{AA} + O_{AA}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)^2 \alpha N - \beta_1 H_{AA} - \mu_1 H_{AA}, && \text{(9b)} \\
H'_{Aa} &= 2\left(\frac{H_{AA} + O_{AA}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)\left(\frac{U_{aa}}{N} + \frac{H_{Aa} + O_{Aa}}{2N}\right)\alpha N - \beta_1 H_{Aa} - \mu_1 H_{Aa}, && \text{(9c)} \\
O'_{AA} &= \beta_1 H_{AA} - \mu_2 O_{AA}, && \text{(9d)} \\
O'_{Aa} &= \beta_1 H_{Aa} - \mu_2 O_{Aa}. && \text{(9e)}
\end{aligned}
$$

with $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$.

3.3 Positive Invariance

We can show that the system for Case 2 is positively invariant using the same methods as in Case 1, except we must further show that $b_{AA}, b_{Aa}, b_{aa} \geq 0$.
Again, assume that $t_0 = 0$, $\hat{X}(0) \equiv \hat{X}_0 \geq 0$, $\alpha, \beta, \hat{p}, \hat{q} > 0$, and that one of the variables is the first to become negative at some point on its trajectory. In order for this to occur, there must be:
i) a first time, t_c, that it leaves the positive quadrant,
ii) it must cross 0, and
iii) it must cross 0 with a negative derivative.

Suppose that $U_{aa}(t_c) = 0$, then

$$U'_{aa}(t_c) = U'_{aa}|_{U_{aa}=0} = b_{aa}\alpha N = \left(\frac{H_{Aa} + O_{Aa}}{2N}\right)^2 \alpha N.$$

Notice that $\left(\frac{U_{aa}+O_{Aa}}{2N}\right)^2 \geq 0$ for any value of H_{Aa} and since U_{aa} is the first variable to cross outside of the positive quadrant we know $N \geq 0$. Thus we have $U'_{aa}(t_c) \geq 0$ which contradicts (iii).
Further, the same analysis can be applied to show that H_{AA} cannot be the first variable to cross outside of the positive quadrant.
If H_{Aa} is the first variable to become negative, then $H_{Aa}(t_c) = 0$ and

$$H'_{Aa}(t_c) = H'_{Aa}|_{H_{Aa}=0} = b_{Aa}\alpha N = 2\left(\frac{H_{AA} + O_{AA}}{N} + \frac{O_{Aa}}{2N}\right)\left(\frac{U_{aa}}{N} + \frac{O_{Aa}}{2N}\right)\alpha N.$$

Since H_{Aa} is the first variable to cross 0, we know that all other variables are still non-negative, and $\left(\frac{H_{AA}+O_{AA}}{N} + \frac{O_{Aa}}{2N}\right)\left(\frac{U_{aa}}{N} + \frac{O_{Aa}}{2N}\right) \geq 0$. Then, $H'_{Aa}(t_c) \geq 0$ which contradicts (iii). Showing positivity of O_{AA} and O_{Aa} follow exactly as in Case 1. Therefore, the Case 2 system is positively invariant.

3.4 Rescaling Case 2

To simplify and further analyze Model 9, we rescale the state variables so that they become proportions of the total population. That is, we define $x_1 = \frac{U_{aa}}{N}, x_2 = \frac{H_{AA}}{N}, x_3 = \frac{H_{Aa}}{N}, x_4 = \frac{O_{AA}}{N}, x_5 = \frac{O_{Aa}}{N}$, and let $a = \frac{\alpha}{\mu_2-\mu_1}, b = \frac{\beta}{\mu_2-\mu_1}$. Notice that a and b are both positive since we assume that obese individuals die at a faster rate than non-obese individuals ($\mu_2 > \mu_1$).

In this case, the genotype frequencies are changed (after scaling) from Expression (8) into:

$$\begin{aligned}
b_{AA} &= p^2 = \left(x_2 + x_4 + \frac{x_3}{2} + \frac{x_5}{2}\right)^2, \\
b_{Aa} &= 2pq = 2\left(x_2 + x_4 + \frac{x_3}{2} + \frac{x_5}{2}\right)\left(x_1 + \frac{x_3}{2} + \frac{x_5}{2}\right), \\
b_{aa} &= q^2 = \left(x_1 + \frac{x_3}{2} + \frac{x_5}{2}\right)^2 \\
&= \left(1 - \left(x_2 + x_4 + \frac{x_3}{2} + \frac{x_5}{2}\right)\right)^2.
\end{aligned}$$

This leads to our new rescaled model:

$$\begin{aligned}
\dot{x_1} &= x_1\left(x_4+x_5-a\right)+\left(x_1+\frac{x_3}{2}+\frac{x_5}{2}\right)^2 a, && \text{(10a)}\\
\dot{x_2} &= x_2\left(x_4+x_5-(b+a)\right)+\left(x_2+x_4+\frac{x_3}{2}+\frac{x_5}{2}\right)^2 a, && \text{(10b)}\\
\dot{x_3} &= x_3(x_4+x_5-(b+a))+2\left(x_2+x_4+\frac{x_3}{2}+\frac{x_5}{2}\right)\left(x_1+\frac{x_3}{2}+\frac{x_5}{2}\right)a, && \text{(10c)}\\
\dot{x_4} &= x_4(x_4+x_5-(1+a))+bx_2, && \text{(10d)}\\
\dot{x_5} &= x_5(x_4+x_5-(1+a))+bx_3, && \text{(10e)}
\end{aligned}$$

where $x_1+x_2+x_3+x_4+x_5=1$. Since we used proportions, we let $x_1=1-(x_2+x_3+x_4+x_5)$. Therefore Model 10 becomes:

$$\begin{aligned}
\dot{x_2} &= x_2(x_4+x_5-(b+a))+\left(x_2+x_4+\frac{x_3}{2}+\frac{x_5}{2}\right)^2 a, && \text{(11a)}\\
\dot{x_3} &= x_3(x_4+x_5-(b+a))+2\left(x_2+x_4+\frac{x_3}{2}+\frac{x_5}{2}\right)\left(1-x_2-x_4-\frac{x_3}{2}-\frac{x_5}{2}\right)a, && \text{(11b)}\\
\dot{x_4} &= x_4(x_4+x_5-(1+a))+bx_2, && \text{(11c)}\\
\dot{x_5} &= x_5(x_4+x_5-(1+a))+bx_3. && \text{(11d)}
\end{aligned}$$

3.4.1 Equilibrium Points

We analyze System 11 by first finding all the equilibrium points $(x_2^*, x_3^*, x_4^*, x_5^*)$. That is, we set the equations on System 11 equal to zero. This yields two equilibria:

Equilibrium	**Value**	**Existence and Stability**
Only Unhealthy Homozygous Recessive	E_1	Always locally asymptotically stable if $\mu_2 > \mu_1$
Only Homozygous Dominant	E_2	Have to use Central Manifold Theorem

E_2^* suggests only the dominant genotype AA for healthy and obese individuals exists. This would mean that over time, the recessive gene breeds out of the population. In order for this point to be biologically reasonable, all of the variables have to be positive and our discriminant $((1+a+b)^2-4b)$ has to be nonnegative. Proving the discriminant is nonnegative, we ended up with $(1-a-b)^2+4a \geq 0$. This inequality always holds.

Also, we have to check that $\Phi = \frac{1}{2}\left(1-a-b+\sqrt{(1+a+b)^2-4b}\right) \geq 0$.

Case 1: $a+b<1$

If $a+b<1$, then $1-a-b>0$ and thus $\Phi>0$.

Case 2: $a + b > 1$

Let $a + b > 1$, then is always true that

$$\begin{aligned} 1 - a - b + \sqrt{(1 + a + b)^2 - 4b} &= 1 - a - b + \sqrt{(1 - a - b)^2 + 4a}, \\ &= 1 - a - b + \sqrt{(-1 + a + b)^2 + 4a}. \end{aligned}$$

Now it is true that

$$\begin{aligned} (-1 + a + b)^2 + 4a &> (a + b - 1)^2, \\ \sqrt{(-1 + a + b)^2 + 4a} &> a + b - 1 \text{ (since a+b -1>0)}, \\ 1 - a - b + \sqrt{(-1 + a + b)^2 + 4a} &> 0. \end{aligned}$$

Case 3: $a + b = 1$

If $a + b = 1$, then

$$\begin{aligned} 1 - a - b + \sqrt{(1 - a - b)^2 + 4a} &= 0 + \sqrt{0^2 + 4a}, \\ 2\sqrt{a} &> 0. \end{aligned}$$

Thus Φ is always positive.

Lastly, we have to prove $\frac{1}{2}\left(1 + a + b - \sqrt{(1 + a + b)^2 - 4b}\right) > 0$. However, it is always true that:

$$\begin{aligned} (1 + a + b)^2 &> (1 + a + b)^2 - 4b, \\ 1 + a + b &> \sqrt{(1 + a + b)^2 - 4b}, \end{aligned}$$

Therefore, $1 + a + b - \sqrt{(1 + a + b)^2 - 4b} > 0$.

Thus, E_2^* is a feasible equilibrium point. In order to analyze the stability of the equilibrium points we calculated the Jacobian matrix at each equilibrium.

Theorem 2 *If $\mu_2 > \mu_1$, then the only homozygous recessive equilibrium is locally asymptotically stable.*

Prove: Stability of E_1. The Jacobian matrix of System 11 at the zero equilibrium is provided by:

$$\mathbf{J} = \begin{pmatrix} -b-a & 0 & 0 & 0 \\ 2a & -b & 2a & a \\ b & 0 & -1-a & 0 \\ 0 & b & 0 & -1-a \end{pmatrix},$$

The eigenvalues for E_1^* are provided by:

$$\begin{aligned} \lambda_1 &= -1-a<0, \\ \lambda_2 &= -a-b<0, \\ \lambda_3 &= \frac{-1}{2}\left(1+a+b+\sqrt{(1+a+b)^2-4b}\right)<0, \\ \lambda_4 &= \frac{-1}{2}\left(1+a+b-\sqrt{(1+a+b)^2-4b}\right)<0. \end{aligned}$$

In order for the point to be stable, all of the eigenvalues should have negative real parts. The first three eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are all negative. While λ_4 is negative if $1+a+b > \sqrt{(1+a+b)^2-4b}$. This was proven in the Equation 12. Therefore, the equilibrium point (1,0,0,0,0) is locally asymptotically stable.

3.4.2 Stability Analysis for E_2^*

In order to calculate the eigenvalues to study the stability of the equilibrium point E_2^*, we use the program Mathematica and obtain:

$$\begin{aligned} \lambda_1 &= -\sqrt{(1+a+b)^2-4b}, \\ \lambda_2 &= \frac{1}{2}(1-a+b-\sqrt{(1+a+b)^2-4b}), \\ \lambda_3 &= \frac{1}{2}\left(-\sqrt{(1+a+b)^2-4b}-\sqrt{(1+a+b)^2-4b}\right) \\ &= -\sqrt{(1+a+b)^2-4b}, \\ \lambda_4 &= \frac{1}{2}\left(-\sqrt{(1+a+b)^2-4b}+\sqrt{(1+a+b)^2-4b}\right)=0. \end{aligned}$$

λ_1 and λ_3 have negative coefficients, therefore they are both negative and $\lambda_4 = 0$.

For λ_2, we need $1-a+b-\sqrt{(1+a+b)^2-4b}<0$. Since we know that $\sqrt{(1+a+b)^2-4b} > 0$, we need $1-a+b<0$, then $a-b>1$, and we will need to use the Center Manifold Theorem to analyze its stability in full.

3.5 Simulations

Simulations for the case if gene factor only with non-constant birth proportions are shown in Figure 4. For these simulations we used the values provided in Table 3.

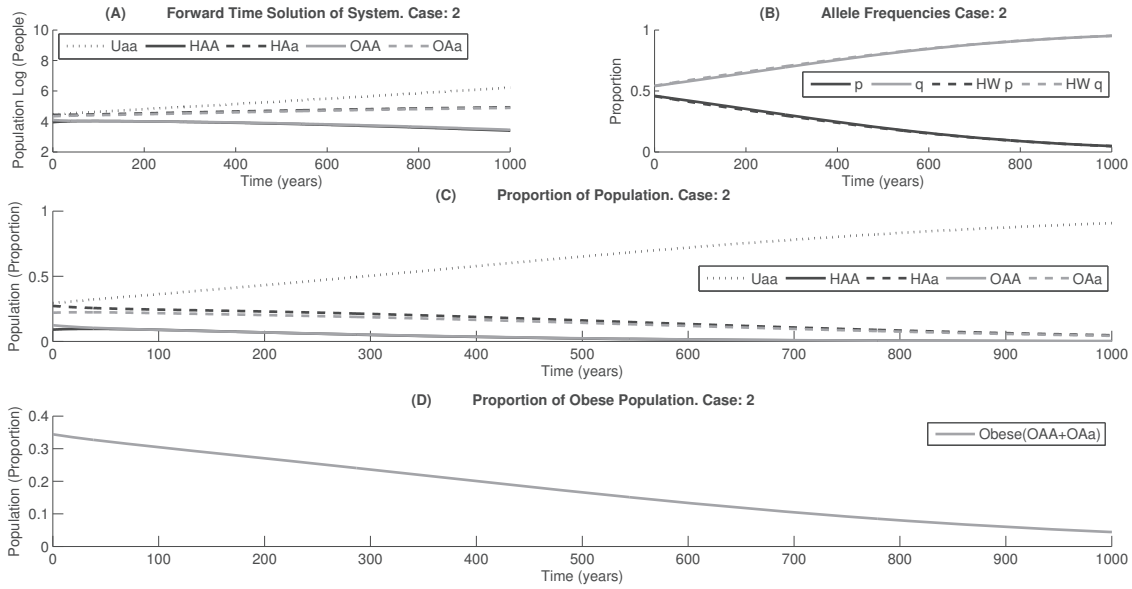

Figure 4: Simulation plots for Case 2.

For Case 2, plot (A) shows that people with homozygous recessive genotype who eat unhealthy increases exponentially, while other populations stay the same or decreases exponentially. Plot (B) shows that the allele frequencies differ from the values predicted by Hardy-Weinberg priciple. Plot (C) shows based on the analysis the homozygous recessive aa goes to one and the rest of the population decreases towards zero. Plot (D) shows that the dominant allele A gets purged out and obese population bred out over more than 1400 years.

4 Scenario 2: Gene and Environment

Under this scenario, we will consider the impact of environmental effect on obesity with genetic factors. We use the same assumptions to simplify the full model (Model 1) as provided in Section 2. Here, by adding the diet factor and comparing this model with the previous model (Model 3), we can have a better understanding of the environmental impact on obesity.

We first study the case (Case 3) where $b_{AA} = \hat{p}^2$, $b_{Aa} = 2\hat{p}\hat{q}$, and $b_{aa} = \hat{q}^2$; that is, the birth proportion of each genotype are constant.

With these assumptions, recall our full model (Model 1) was reduced to Model 2:

$$\begin{aligned}
U'_{aa} &= b_{aa}\alpha N - \mu_1 U_{aa}, \\
H'_{AA} &= b_{AA}\alpha N - \beta_1\left(\hat{X}\right) H_{AA} + \gamma_1\left(\hat{X}\right) O_{AA} - \mu_1 H_{AA}, \\
H'_{AA} &= b_{Aa}\alpha N - \beta_1\left(\hat{X}\right) H_{Aa} + \gamma_1\left(\hat{X}\right) O_{Aa} - \mu_1 H_{Aa}, \\
O'_{AA} &= \beta_1\left(\hat{X}\right) H_{AA} - \gamma_1\left(\hat{X}\right) O_{AA} - \mu_2 O_{AA}, \\
O'_{Aa} &= \beta_1\left(\hat{X}\right) H_{Aa} - \gamma_1\left(\hat{X}\right) O_{Aa} - \mu_2 O_{Aa},
\end{aligned}$$

where $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$ and $\hat{X} = [U_{aa}, H_{AA}, H_{Aa}, O_{AA}, O_{Aa}]$.

For Cases 3 and 4 we define $\beta_1\left(\hat{X}\right) = \hat{\beta}\left(\frac{U_{aa}+O_{AA}+O_{Aa}}{N}\right)$ and $\gamma_1\left(\hat{X}\right) = \hat{\gamma}\left(\frac{H_{AA}+H_{Aa}}{N}\right)$.

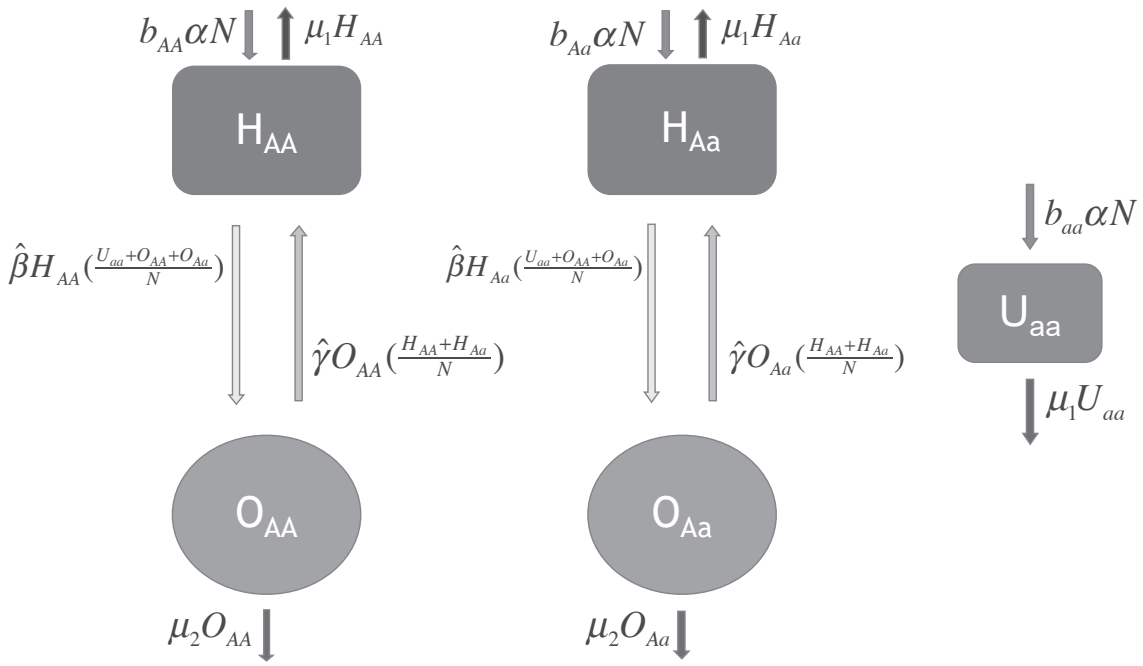

Figure 5: Flow chart of reduced model with gene and environment factors (nonlinear β).

We consider healthy individuals to be recruited into obesity by the effect of the environment, specifically by interactions with unhealthy and obese individuals (Figure 5). We consider obese individuals to be recruited into the healthy class by the effect of the environment as well, but by interacting with healthy individuals.

4.1 Simulations for Constant Birth Ratios

The purpose of this case (Case 3) is to find values for $\hat{\beta}$ and $\hat{\gamma}$ to have the Obese proportion be 30% to match the current data. Here we calculated $\hat{\beta}$ and $\hat{\gamma}$ as provided in Table 4 based on the graph below.

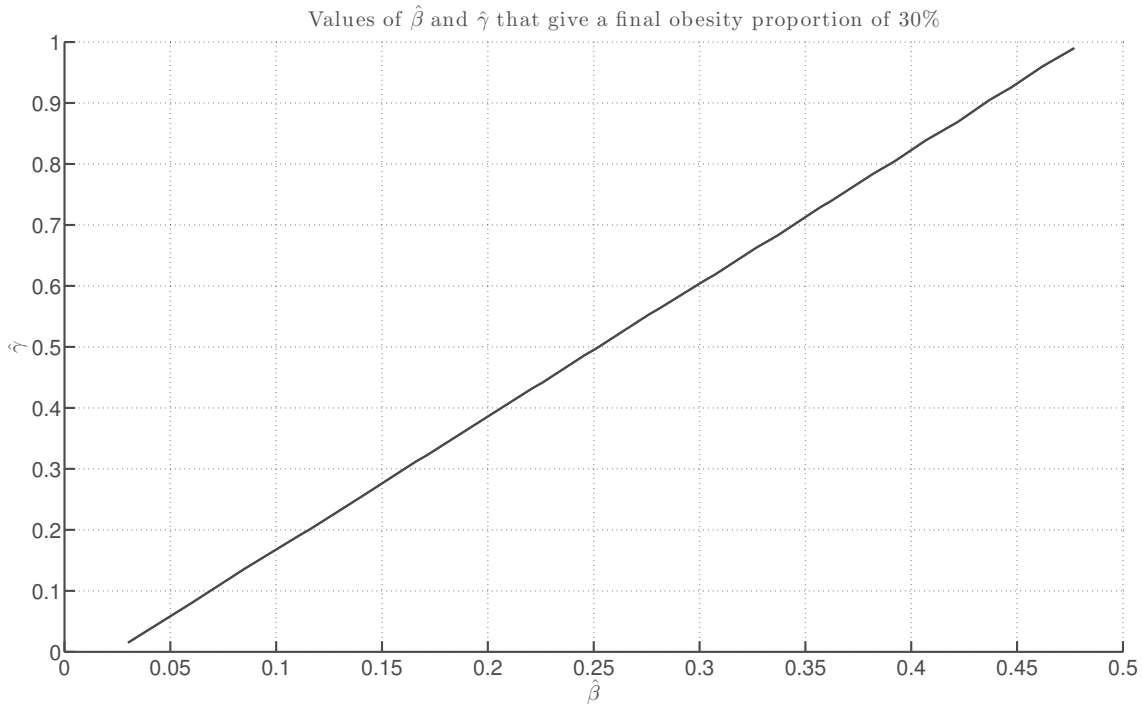

Figure 6: Different combinations of β and γ that make the final obese population 30%.

Figure 6 shows the possible β and γ values that make the final obese proportion 30%. This is to match data from the CDC [5]. We ran a combination of different β and γ values that would yield the final population to match the data.

Table 4. $\hat{\beta}$ and $\hat{\gamma}$ for Case 3 Simulation

Parameter	Meaning	Calculated Value
$\hat{\beta}$	linear transfer function from healthy eaters to obesity with genotype AA or Aa	0.0603
$\hat{\gamma}$	recovery function from obesity to healthy genotypes AA andAa $\left(\text{time}^{-1}\right)$	0.0804

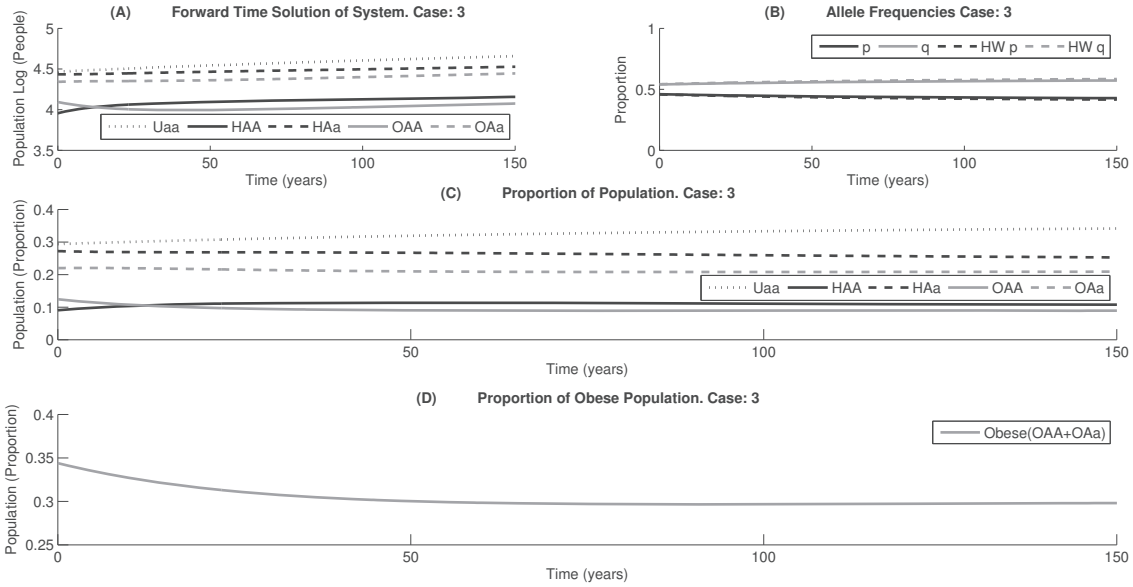

Figure 7: Simulation plots for Case 3.

In this case, plots in Figure 7 show that while the total population increases, obese population decreases at first and levels off at approximately 30% in about 250 years. The recessive allele a outcompetes the dominant allele A.

4.2 Analysis for Non-Constant Birth Ratios

For Case 4, the birth proportions, that is b_{AA}, b_{Aa}, and b_{aa}, are defined in Equations 8 from Section 3.2. Thus, plugging these equations into Model 2, the simplified model becomes

$$U'_{aa} = \left(\frac{U_{aa}}{N} + \frac{H_{Aa}+O_{Aa}}{2N}\right)^2 \alpha N - \mu_1 U_{aa}, \tag{4a}$$

$$H'_{AA} = \left(\frac{H_{AA}+O_{AA}}{N} + \frac{H_{Aa}+O_{Aa}}{2N}\right)^2 \alpha N - \beta_1\left(\hat{X}\right) H_{AA} + \gamma_1\left(\hat{X}\right) O_{AA} - \mu_1 H_{AA}, \tag{4a}$$

$$H'_{Aa} = 2\left(\frac{H_{AA}+O_{AA}}{N} + \frac{H_{Aa}+O_{Aa}}{2N}\right)\left(\frac{U_{aa}}{N} + \frac{H_{Aa}+O_{Aa}}{2N}\right)\alpha N - \beta_1\left(\hat{X}\right) H_{Aa} + \gamma_1\left(\hat{X}\right) O_{Aa} - \mu_1 H_{Aa}, \tag{4a}$$

$$O'_{AA} = \beta_1\left(\hat{X}\right) H_{AA} - \gamma_1\left(\hat{X}\right) O_{AA} - \mu_2 O_{AA}, \tag{4b}$$

$$O'_{Aa} = \beta_1\left(\hat{X}\right) H_{Aa} - \gamma_1\left(\hat{X}\right) O_{Aa} - \mu_2 O_{Aa}, \tag{4c}$$

where $N = U_{aa} + H_{AA} + H_{Aa} + O_{AA} + O_{Aa}$ is non-constant, since the population changes

over time.

Substitution of our variables that are scaled to proportions into Model 4 and remembering $x_1 = 1 - (x_2 + x_3 + x_4 + x_5)$, we get the rescaled system

$$\begin{aligned}
\dot{x_2} &= \left(x_2\left(x_2 + 2x_4 + x_3 + x_5\right) + x_4\left(x_4 + x_3 + x_5\right) + \left(\frac{x_3 + x_5}{2}\right)^2 \right) a && \text{(5a)}\\
&\quad + c(x_2 + x_3)x_4 + \left(-b + b(x_2 + x_3) - a + x_4 + x_5\right) x_2, && \text{(5b)}\\
\dot{x_3} &= \left(2\left(x_2 + x_4 + \frac{x_3 + x_5}{2} \right)\left(1 - \left(x_2 + x_4 + \frac{x_3 + x_5}{2} \right) \right) \right) a && \text{(5c)}\\
&\quad + c(x_2 + x_3)x_5 + \left(-b + b(x_2 + x_3) - a + x_4 + x_5\right) x_3, && \text{(5d)}\\
\dot{x_4} &= b(1 - x_2 - x_3)x_2 + \left(-c(x_2 + x_3) - a - 1 + x_4 + x_5\right) x_4, && \text{(5e)}\\
\dot{x_5} &= b(1 - x_2 - x_3)x_3 + \left(-c(x_2 + x_3) - a - 1 + x_4 + x_5\right) x_5, && \text{(5f)}
\end{aligned}$$

where $c = \dfrac{\gamma}{\mu_2 - \mu_1} > 0$.

4.2.1 Equilibria

Setting the Equations of System 5 equal to zero provided the equilibrium points of the system $(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*)$.

Equilibrium	**Value**	**Existence and Stability**
Only Unhealthy Homozygous Recessive	E_1	Always locally asymptotically stable if $\mu_2 > \mu_1$
Only Healthy Homozygous Dominant	E_2	Unstable if E_3 exists
Only Homozygous Dominant	E_3	Exists if $b > 1 + a + c$

Biologically, E_1 is only the homozygous recessive population U_{aa}. E_2 is only the homozygous dominant healthy population H_{AA}. E_3 is the homozygous dominant populations H_{AA} and O_{AA}. E_3 only exists if $b > 1 + a + c$.

4.2.2 Stability of Equilibria

From Mathematica, the eigenvalues of the homozygous recessive equilibrium (E_1) are

$\lambda_1 = -(1+a)$, $\lambda_2 = -(a+b)$, $\lambda_3 = \frac{-(1+a+b)-\sqrt{(1+a+b)^2-4b}}{2}$, $\lambda_4 = \frac{-(1+a+b)+\sqrt{(1+a+b)^2-4b}}{2}$.

From Section 3.4.1, the only homozygous recessive equilibrium is locally stable only if $\mu_2 > \mu_1$.

The eigenvalues from Mathematica for E_2 are

$\lambda_1 = 0$, $\lambda_2 = -a$, $\lambda_3 = -(1 + a + c)$, $\lambda_4 = b - (1 + a + c)$.

The eigenvalues for E_3 are

$\lambda_1 = 0$, $\lambda_2 = 1 + a + c - b$, $\lambda_3 = 1 + \frac{a}{1+c-b}$, $\lambda_4 = 1 + \frac{a}{1+c-b} - b$.
If E_3 exists then, E_2 is unstable. Assuming $\mu_2 > \mu_1$ and $b > 1 + a + c > 1 + c$, we want to show that the non-zero eigenvalues of E_3 are negative, then

$$1 + \frac{a}{1+c-b} < 0 \Rightarrow \frac{a}{b-(1+c)} > 1 \Rightarrow a > b - (1+g) \Rightarrow b < 1 + a + g,$$

which contradicts the existence of E_3, so E_3 is unstable.

If E_3 does not exist, E_2 has three negative eigenvalues and one eigenvalue that is zero. The Central Manifold Theorem is needed to further analyze for stability.

4.2.3 Simulations for Case 4: Non-Fixed Birth Proportions, Non-Constant β_1

The simulations for Case 4 are shown in Figure 8. For these simulations we use the parameters provided in Table 3.

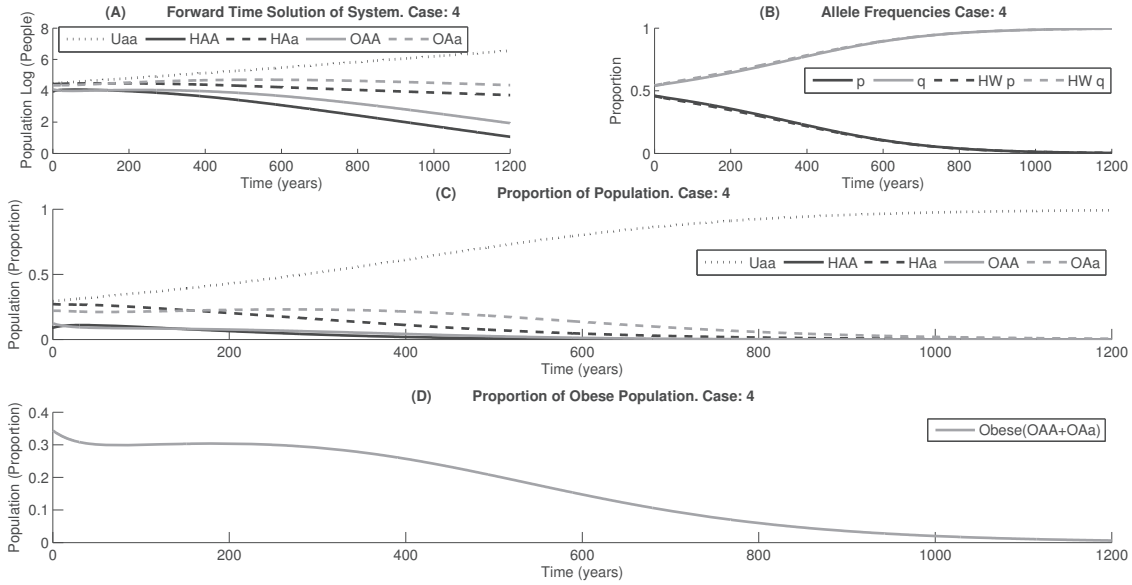

Figure 8: Simulation plots for Case 4.

Plot (A) shows that proportions of populations are decreasing except for the homozygous recessive alleles. People who have the homozygous recessive genotype and eat unhealthy totally outbreed obese populations after 1000 years.

5 Results

In this section, we discuss and further compare the scenarios presented in this work.

5.1 Constant vs non-constant genotype frequencies

In this section we study the impact of the genotype frequencies on obesity. We compare the two cases, with constant vs non-constant genotype frequencies, in each of the scenarios to understand what impact genetic inheritance plays in the dynamics of obesity progression. Comparing case 1 (where we only consider genetics factor and the birth ratios are constant) versus case 2 (where the birth ratios are not constant), we noticed that the allele frequencies under case 1 stabilize faster than that of case 2. Additionally, the population with the homozygous recessive genotype aa has the largest population size at equilibrium. The proportion of each allele differs very slightly from proportions predicted by Hardy-Weinberg Principle. In contrast, for case 2, when the genotype frequencies depends on the state variable, the proportion of the obese population goes to zero and the proportion of the population with the homozygous recessive genotype asymptotically approaches one. Computationally, that is it approaches the (1,0,0,0,0) equilibria.

For scenario 2, the differences in behavior, for the cases with environmental factors, between case 3 (where the birth proportions are constant) and 4 (where the birth proportions are not constant) are essentially similar as in scenario 1. This means that when the genotype frequencies are constant, the proportion of each class reaches a coexistence equilibrium quickly; but, when the genotype frequency is birth dependent, the proportion of the population without the obesity allele variant will grow asymptotically to 1. Also, in case 3, the proportion of unhealthy people without the gene variant will surpass the proportion of obese population faster in the short term than in case 4. The population of the unhealthy eaters without the gene.

5.2 Genetics factor vs genetics and environment factor

To see if the environment factor plays a significant role in obesity, we compared the cases with and without environmental factors, that is case 1 to case 3, and case 2 to case 4. In case 1, we chose birth and death rates according to data and selected the rate of transfer (β) to the obese class to give us a final proportion of obese individuals that roughly matches the CDC data on obesity in the United States (34.9%). We also chose a combination of β and γ for case 3 that gave us the same final proportion (approximately 30%) of obese individuals for the year 2015. The inclusion of environmental effects and the capacity for people to recover from obesity means that the social-influenced transfer to obesity requires a higher transfer rate (β) than in case 1, which only considered the genetics factor.

While comparing cases 1 and 3, we chose different β values to analyze how this parameter

impacts the model. We assumed β is between 0 and 1 and we arbitrarily choose $\beta = 0.2$ and $\beta = 0.8$. The reason that we decided on these two values was to see the impact of β in the extremes cases (i.e. an average of 5 years waiting time to become obese versus an average 1.25 year waiting time). Simulations on Figures 3 and 7 shows that β increases the proportion of obese people but does not change the qualitative dynamics of the model. Additionally, we noticed that in case 1 (see Figure 3), the heterozygous dominant obese class is initially greater than the homozygous recessive unhealthy population, but the homozygous recessive genotype surpasses it after approximately 150 years. In contrast, the heterozygous obese class always surpasses the homozygous recessive genotype. So, the obese population is higher than the population of non-obese unhealthy eaters.

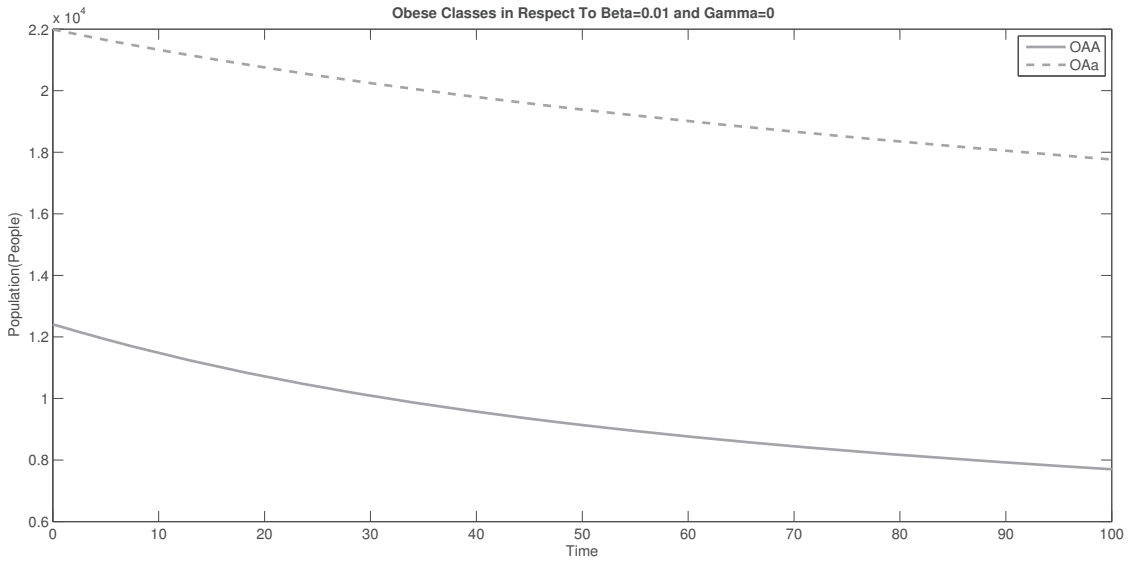

Figure 9: Simulation plots of case 2 with genetics factor, and non-constant birth proportions. $\beta = 0.01$, $\gamma = 0$.

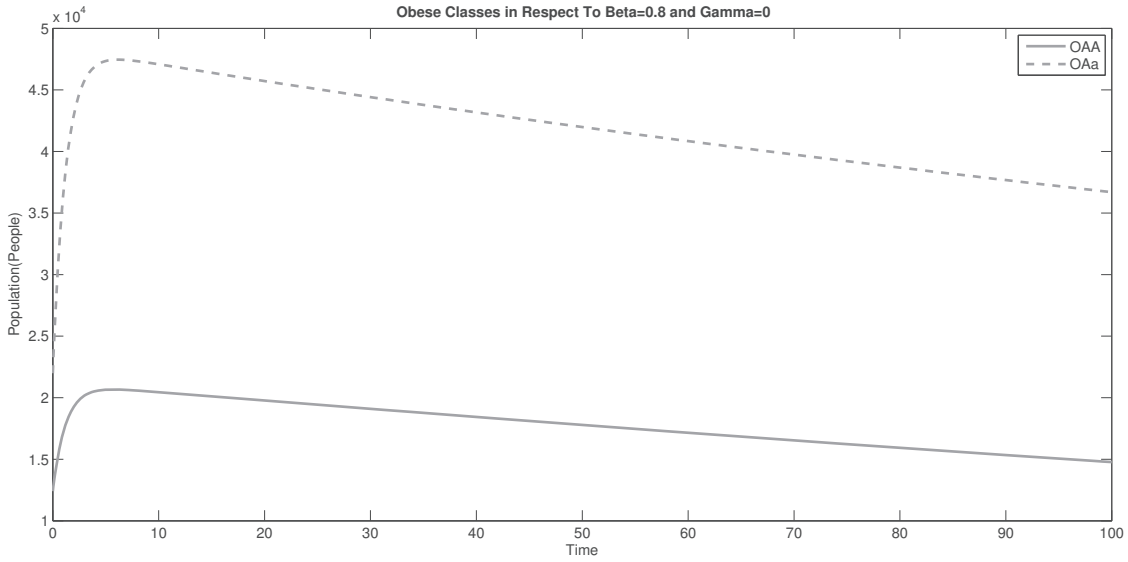

Figure 10: Simulation plots of case 2 with genetics factor, and non-constant birth proportions. $\beta = 0.8$, $\gamma = 0$.

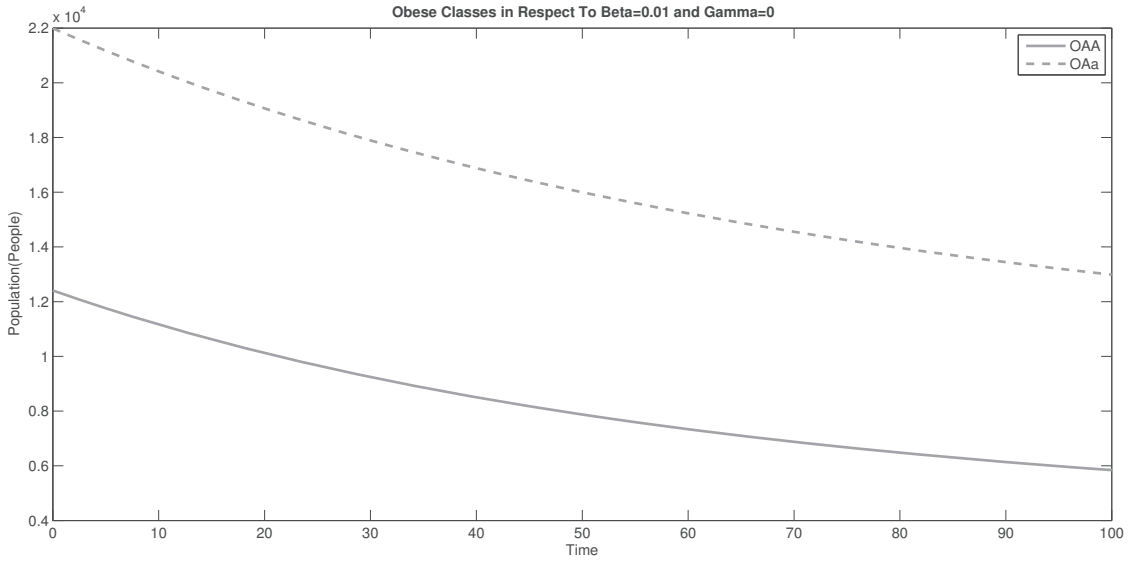

Figure 11: Simulation plots of case 4 with both genetics and environment factors, and non-constant birth proportions. $\beta = 0.01$, $\gamma = 0.8$.

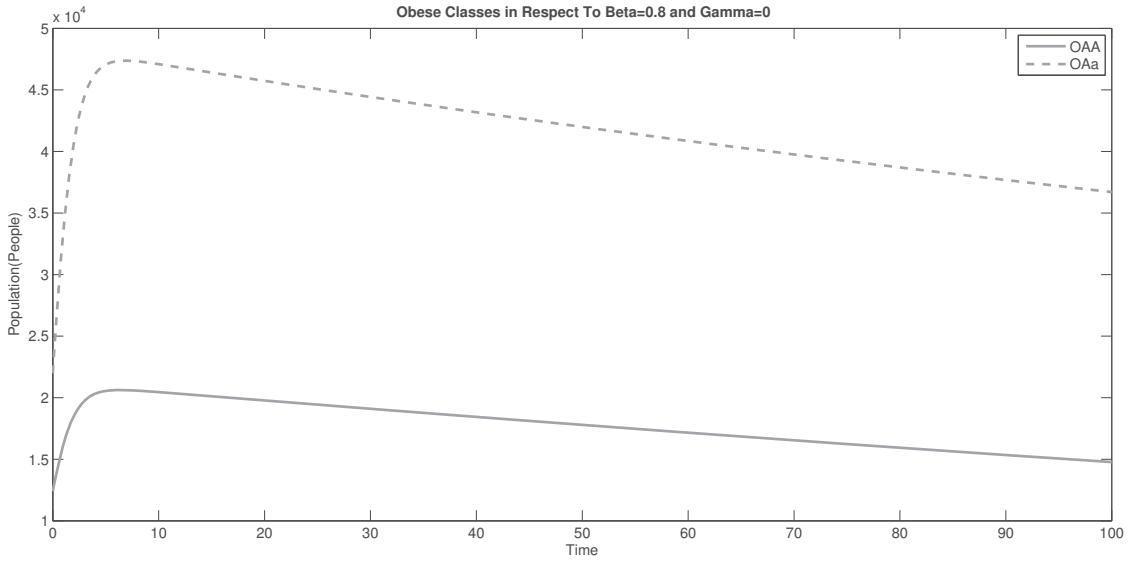

Figure 12: Simulation plots of case 4 with both genetics and environment factors, and non-constant birth proportions. $\beta = 0.8$, $\gamma = 0.8$.

When we compare case 2 and case 4 (case 4 adds an environment factor to the case 2, and both have non-constant birth proportions), we noticed that the graphs are qualitatively identical see Figures (9-12). The transient dynamics vary slightly, but it seems very clear that the environmental factor does not play a large role in the long term dynamics of obesity.

The obesity transfer rate (β) strongly affects the proportion of people in the obese population, especially when birth proportions are fixed. As β increases, the obese population naturally increases. Finally, environmental factors only seem to play a roll in obesity when the genotype frequency dependent birth proportions are constant, but when they depend on the state variables, the environment factor does not affect obesity.

6 Conclusion

From these results, we conclude that when the birth proportions are forced constant, it allows the proportions of populations to reach a coexistence equilibrium. Under case 3, the proportion of obese populations reach equilibrium faster than under case 1. With the same constant birth proportions, the difference suggests that environment factor actually plays a role in our model. However, because case 1 and 3 are biological infeasible in the real world (birth proportions cannot be constant), we further compare the results of case 2 and 4, where we consider non-constant birth proportions, and case 4 adds an environment factor to

case 2). Results from stability analysis and simulations for these two cases match. The only feasible equilibrium for case 2 and 4 is when the homozygous recessive population outbreeds the populations with the dominant allele, including the obese populations. Therefore, the effect of environmental factors, especially diet, on the dynamics of obesity are negliglble at best. Fitness and genetic selection trumps any environmental bias in the long term. But in the short term, unhealthy diet still greatly increases obesity populations.

7 Future Work

In this work, we constructed simplified models to study the joint effects of genetics and environment. In reality, birth proportions cannot be constant, and there is not just one pathway that leads to obesity. Thus in the future, we will expand our simplified model to consider more complicated and realistic cases. For example, we will incorporate the obesity function $\rho(\bar{X})$ and recover function $\xi(\bar{X})$ into the model. Also, we will allow heterozygous recessive genotype aa to become an obesity-inducing allele. In additon, we will construct a stochastic model that considers mutations, random drifts and other random processes that may greatly change allele frequencies under stochastic conditions. This would make our model more realistic. Lastly, a sensitivity analysis on important parameters, such as β and γ, should be conducted to show their effects on the equlibria, and determine their relative importance over time.

8 Acknowledgements

We would like to thank Dr. Carlos Castillo-Chavez, Executive Director of the Mathematical and Theoretical Biology Institute (MTBI), for giving us this opportunity to participate in this research program. We would also like to thank Summer Director Dr. Anuj Mubayi, coordinators Preston Swan and Ciera Duran for their efforts in planning and executing the day-to-day activities of MTBI. In addition, we thank Robert Mamada for his support. This research was conducted in MTBI at the Simon A. Levin Mathematical, Computational and Modeling Sciences Center (SAL MCMSC) at Arizona State University (ASU). This project has been partially supported by grants from the National Science Foundation (DMS 1263374), the National Security Agency (H98230-15-1-0021), the Office of the President of ASU, and the Office of the Provost at ASU.

References

[1] Kenneth F Adams, Arthur Schatzkin, Tamara B Harris, Victor Kipnis, Traci Mouw, Rachel Ballard-Barbash, Albert Hollenbeck, and Michael F Leitzmann. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *New England Journal of Medicine*, 355(8):763–778, 2006.

[2] Tariq Ahmad, I-Min Lee, Guillaume Paré, Daniel I Chasman, Lynda Rose, Paul M Ridker, and Samia Mora. Lifestyle interaction with fat mass and obesity-associated (fto) genotype and risk of obesity in apparently healthy us women. *Diabetes care*, 34(3):675–680, 2011.

[3] Fred Brauer and Christopher Kribs. An introduction to dynamics systems for biological modeling. 2015.

[4] CDC. Deaths and mortality, 2013. URL: `http://www.cdc.gov/nchs/fastats/deaths.htm`.

[5] CDC. Obesity prevalence in 2013 varies across states and territories, 2013. URL: `http://www.cdc.gov/obesity/data/prevalence-maps.html`.

[6] James F Crow, Motoo Kimura, et al. An introduction to population genetics theory. *An introduction to population genetics theory.*, 1970.

[7] Keisuke Ejima, Kazuyuki Aihara, and Hiroshi Nishiura. Modeling the obesity epidemic: social contagion and its implications for control. *Theor Biol Med Model*, 10(1):17, 2013.

[8] Parvez Hossain, Bisher Kawar, and Meguid El Nahas. Obesity and diabetes in the developing world—a growing challenge. *New England journal of medicine*, 356(3):213–215, 2007.

[9] Gordon S Huggins, Samantha Berger, and Jeanne M McCaffery. Can genetics modify the influence of healthy lifestyle on lipids in the context of obesity and type 2 diabetes? *Current Cardiovascular Risk Reports*, 9(7):1–7, 2015.

[10] Kihwa Kang, Cuilin Zhang, Peter Kraft, Lu Qi, David G Hunter, Rob M van Dam, Chih-Hao Lee, and Frank B Hu. Fat mass–and obesity-associated (fto) gene variant is associated with obesity. 2008.

[11] Sihua Peng, Yimin Zhu, Fangying Xu, Xiaobin Ren, Xiaobo Li, and Maode Lai. Fto gene polymorphisms and obesity risk: a meta-analysis. *BMC medicine*, 9(1):71, 2011.

[12] Emily Sonestedt, Charlotta Roos, Bo Gullberg, Ulrika Ericson, Elisabet Wirfält, and Marju Orho-Melander. Fat and carbohydrate intake modify the association between genetic variation in the fto genotype and obesity. *The American journal of clinical nutrition*, 90(5):1418–1425, 2009.

[13] Willer C.J. Berndt S.I. Monda K.L. Thorleifsson G. Jackson A.U. Allen H.L. Lindgren C.M. Luan J. Magi R. Speliotes, E.K. Giant consortium data files@ONLINE, 2013. URL: `https://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files`.

[14] Ba Swinburn, Seidell JC, James WPT, et al. Diet, nutrition and the prevention of excess weight gain and obesity. *Public health nutrition*, 7(1a):123–146, 2004.

[15] Diana M Thomas, Marion Weedermann, Bernard F Fuemmeler, Corby K Martin, Nikhil V Dhurandhar, Carl Bredlau, Steven B Heymsfield, Eric Ravussin, and Claude Bouchard. Dynamic model predicting overweight, obesity, and extreme obesity prevalence trends. *Obesity*, 22(2):590–597, 2014.

[16] Esther Zimmermann, Sofia II Kring, Tina L Berentzen, Claus Holst, Tune H Pers, Torben Hansen, Oluf Pedersen, Thorkild IA Sørensen, and Tine Jess. Fatness-associated fto gene variant increases mortality independent of fatness–in cohorts of danish men. 2009.

9 Appendix

Here we show the details of the calculation of stabilities for various equilibrium points.

9.1 Case 1

The Jacobian matrix of System 3.1.2 at the extinction equilibria $(0,0,0,0,0)$ is provided by:

$$\det\left(\mathbf{J}-\lambda\mathbf{I}\right)=\det\begin{pmatrix} q^2\alpha-\mu_1-\lambda & q^2\alpha & q^2\alpha & q^2\alpha & q^2\alpha \\ p^2\alpha & p^2\alpha-\beta-\mu_1-\lambda & p^2\alpha & p^2\alpha & p^2\alpha \\ 2pq\alpha & 2pq\alpha & 2pq\alpha-\beta-\mu_1-\lambda & 2pq\alpha & 2pq\alpha \\ 0 & \beta & 0 & -\mu_2-\lambda & 0 \\ 0 & 0 & \beta & 0 & -\mu_2-\lambda \end{pmatrix},$$

Expanding this out yields

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right)=&+\left(-\beta-\mu_1-\lambda\right)\left(-\beta-\mu_1-\lambda\right)\left(q^2\alpha\right)\left(-\mu_2-\lambda\right)\left(-\mu_2-\lambda\right)\\ &+\left(-\mu_1-\lambda\right)\left(-\beta-\mu_1-\lambda\right)\left(-\mu_2-\lambda\right)\left(p^2\alpha\right)\left(-\beta-\mu_2-\lambda\right)\\ &+\left(-\mu_1-\lambda\right)\left(-\beta-\mu_1-\lambda\right)\left(-\mu_2-\lambda\right)\left(\left(2pq\alpha-\beta-\mu_1-\lambda\right)\left(-\mu_2-\lambda\right)-\left(2pq\alpha\right)\left(\beta\right)\right).\end{aligned}$$

Factoring out $\left(\beta+\mu_1+\lambda\right)\left(\mu_2+\lambda\right)$ and after simplification,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_2+\lambda\right)\\ &\quad\left[\left(\beta+\mu_1+\lambda\right)\left(q^2\alpha\right)\left(\mu_2+\lambda\right)+\left(\mu_1+\lambda\right)\left(p^2\alpha\right)\left(\beta+\mu_2+\lambda\right)\right.\\ &\quad\left.+\left(\mu_1+\lambda\right)\left(2pq\alpha\left(\mu_2+\lambda\right)-\left(\beta+\mu_1+\lambda\right)\left(\mu_2+\lambda\right)+2pq\alpha\beta\right)\right].\end{aligned}$$

Factoring out alpha and after grouping,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_2+\lambda\right)\\ &\quad\left[\left(\mu_1+\lambda\right)\left(\alpha\left(q^2+p^2+2pq\right)-\beta-\mu_1-\lambda\right)\left(\mu_2+\lambda\right)+\left(\mu_1+\lambda\right)\alpha\left(p^2+2pq\right)\left(\beta\right)\right.\\ &\quad\left.+\beta q^2\alpha\left(\mu_2+\lambda\right)\right]\end{aligned}$$

Realizing $p^2+2pq+q^2=1$ since $p+q=1$,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_2+\lambda\right)\\ &\quad\left[\left(\mu_1+\lambda\right)\left(\alpha-\beta-\mu_1-\lambda\right)\left(\mu_2+\lambda\right)+\beta\left(\mu_1+\lambda\right)\alpha\left(p^2+2pq\right)+\beta q^2\alpha\left(\mu_2+\lambda\right)\right]\end{aligned}$$

Since $p^2+2pq+q^2=1$, $p^2+2pq=1-q^2$, so we can get rid of p so our equations look like

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_2+\lambda\right)\\ &\quad \left[\left(\mu_1+\lambda\right)\left(\alpha-\beta-\mu_1-\lambda\right)\left(\mu_2+\lambda\right)+\beta\left(\mu_1+\lambda\right)\alpha\left(1-q^2\right)+\beta q^2\alpha\left(\mu_2+\lambda\right)\right]\end{aligned}$$

But we know $\mu_2 > \mu_1$ and they are both constants, so let's write μ_2 in terms of μ_1

$$\mu_2 = \mu_1 + \zeta$$

plugging this in,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_1+\zeta+\lambda\right)\\ &\quad \left[\left(\mu_1+\lambda\right)\left(\alpha-\beta-\mu_1-\lambda\right)\left(\mu_1+\zeta+\lambda\right)+\beta\left(\mu_1+\lambda\right)\alpha\left(1-q^2\right)\right.\\ &\quad \left.+\beta q^2\alpha\left(\mu_1+\zeta+\lambda\right)\right]\end{aligned}$$

separating out the ζ terms from the body,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_1+\zeta+\lambda\right)\\ &\quad \left[\left(\mu_1+\lambda\right)\left(\alpha-\beta-\mu_1-\lambda\right)\left(\mu_1+\lambda\right)+\left(\mu_1+\lambda\right)\left(\alpha-\beta-\mu_1-\lambda\right)\left(\zeta\right)\right.\\ &\quad \left.+\beta\left(\mu_1+\lambda\right)\alpha\left(1-q^2\right)+\beta q^2\alpha\left(\mu_1+\lambda\right)+\beta q^2\alpha\left(\zeta\right)\right]\end{aligned}$$

Grouping the $\mu_1+\lambda$,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_1+\zeta+\lambda\right)\\ &\quad \left[\left(\mu_1+\lambda\right)^2\left(\alpha-\beta-\mu_1-\lambda\right)+\left(\mu_1+\lambda\right)\left(\alpha-\beta-\mu_1-\lambda\right)\left(\zeta\right)+\beta\left(\mu_1+\lambda\right)\alpha+\beta q^2\alpha\left(\zeta\right)\right]\end{aligned}$$

grouping again,

$$\begin{aligned}\det\left(\mathbf{J}-\lambda\mathbf{I}\right) &= \left(\beta+\mu_1+\lambda\right)\left(\mu_1+\zeta+\lambda\right)\\ &\quad \left[\left(\mu_1+\lambda\right)^2\left(\alpha-\beta-\mu_1-\lambda\right)+\left(\mu_1+\lambda\right)\left(\zeta\left(\alpha-\beta-\mu_1-\lambda\right)+\beta\alpha\right)+\beta q^2\alpha\zeta\right]=0\end{aligned}$$

This is our final determinant up to this point. Notice both

$$\begin{aligned}\lambda_1 &= -\beta-\mu_1,\\ \lambda_2 &= -\mu_1-\zeta=-\mu_2,\end{aligned}$$

which are both negative.

9.1.1 Finding the values of the other eigenvalues

Removing the two factored eigenvalues,

$$\begin{aligned}
0 &= (\mu_1+\lambda)^2(\alpha-\beta-\mu_1-\lambda)+(\mu_1+\lambda)(\zeta(\alpha-\beta-\mu_1-\lambda)+\beta\alpha)+\beta q^2\alpha\zeta \\
&= (\mu_1+\lambda)^2(\alpha-\beta)+(\mu_1+\lambda)^2(-\mu_1-\lambda)+(\mu_1+\lambda)\zeta(\alpha-\beta-\mu_1-\lambda)+(\mu_1+\lambda)\beta\alpha \\
&\quad +\beta q^2\alpha\zeta \\
&= (\mu_1+\lambda)^2(\alpha-\beta)+(-1)(\mu_1+\lambda)^3+(\mu_1+\lambda)\zeta(\alpha-\beta)+(\mu_1+\lambda)\zeta(-\mu_1-\lambda) \\
&\quad +(\mu_1+\lambda)\beta\alpha+\beta q^2\alpha\zeta \\
&= (\mu_1+\lambda)^2(\alpha-\beta)+(-1)(\mu_1+\lambda)^3+(\mu_1+\lambda)\zeta(\alpha-\beta)+(\mu_1+\lambda)^2(-\zeta) \\
&\quad +(\mu_1+\lambda)\beta\alpha+\beta q^2\alpha\zeta \\
&= (-1)(\mu_1+\lambda)^3+(\mu_1+\lambda)^2(\alpha-\beta-\zeta)+(\mu_1+\lambda)(\zeta(\alpha-\beta)+\beta\alpha)+\beta q^2\alpha\zeta \\
&= (-1)\left(\mu_1^3+3\mu_1^2\lambda+3\mu_1\lambda^2+\lambda^3\right)+\left(\mu_1^2+2\mu_1\lambda+\lambda^2\right)(\alpha-\beta-\zeta)+(\mu_1+\lambda)(\zeta(\alpha-\beta)+\beta\alpha)+\beta q^2\alpha\zeta \\
&= (-1)\lambda^3+\lambda^2(-3\mu_1+\alpha-\beta-\zeta)+\lambda\left(-3\mu_1^2+2\mu_1(\alpha-\beta-\zeta)+(\zeta(\alpha-\beta)+\beta\alpha)\right) \\
&\quad -\mu_1^3-2\mu_1+\mu_1^2(\alpha-\beta-\zeta)+\mu_1(\zeta(\alpha-\beta)+\beta\alpha)+\beta q^2\alpha\zeta \\
&= \lambda^3(-1)+\lambda^2(-3\mu_1+\alpha-\beta-\zeta)+\lambda\left(-3\mu_1^2+2\mu_1(\alpha-\beta-\zeta)+\zeta(\alpha-\beta)+\beta\alpha\right) \\
&\quad +\left(-\mu_1^3+\mu_1^2(\alpha-\beta-\zeta)+\mu_1(\zeta(\alpha-\beta)+\beta\alpha-2)+\beta q^2\alpha\zeta\right) \\
&= \lambda^3(a_1)+\lambda^2(a_2)+\lambda(a_3)+a_4,
\end{aligned}$$

where

$$\begin{aligned}
a_1 &= -1 \\
a_2 &= -3\mu_1+\alpha-\beta-\zeta \\
a_3 &= -3\mu_1^2+2\mu_1(\alpha-\beta-\zeta)+\zeta(\alpha-\beta)+\beta\alpha \\
a_4 &= -\mu_1^3+\mu_1^2(\alpha-\beta-\zeta)+\mu_1(\zeta(\alpha-\beta)+\beta\alpha-2)+\beta q^2\alpha\zeta
\end{aligned}$$

Using Descartes Rule of Signs, let's look at the coefficients of this cubic and set them all to be less than zero so there are no changes in sign. We are trying to prove if $\alpha < \mu_1$, we are stable.
This first coefficient a_1 is always negative,

$$(-1) < 0.$$

Looking at the a_2,

$$\begin{aligned}
(-3\mu_1+\alpha-\beta-\zeta) &< 0, \\
\alpha &< 3\mu_1+\beta+\zeta, \\
\alpha &< \mu_1+2\mu_1+\beta+\zeta,
\end{aligned}$$

and a_3:

$$\begin{aligned}
-3\mu_1^2+2\mu_1\left(\alpha-\beta-\zeta\right)+\zeta\left(\alpha-\beta\right)+\beta\alpha &< 0,\\
2\mu_1\alpha+\zeta\alpha+\beta\alpha &< 3\mu_1^2+2\mu_1\left(\beta+\zeta\right)+\zeta\beta,\\
\alpha\left(2\mu_1+\zeta+\beta\right) &< 3\mu_1^2+2\mu_1\left(\beta+\zeta\right)+\zeta\beta,\\
\alpha &< \frac{3\mu_1^2+2\mu_1\left(\beta+\zeta\right)+\zeta\beta}{\left(2\mu_1+\zeta+\beta\right)},\\
\alpha &< \frac{\mu_1\left(3\mu_1+2\beta+2\zeta\right)+\zeta\beta}{\left(2\mu_1+\zeta+\beta\right)},\\
\alpha &< \frac{\mu_1\left(3\mu_1+2\beta+2\zeta\right)}{\left(2\mu_1+\zeta+\beta\right)}+\frac{\zeta\beta}{\left(2\mu_1+\zeta+\beta\right)},\\
\alpha &< \frac{\mu_1\left(2\mu_1+\beta+\zeta+\mu_1+\beta+\zeta\right)}{\left(2\mu_1+\zeta+\beta\right)}+\frac{\zeta\beta}{\left(2\mu_1+\zeta+\beta\right)},\\
\alpha &< \mu_1+\frac{\mu_1\left(\mu_1+\beta+\zeta\right)}{\left(2\mu_1+\zeta+\beta\right)}+\frac{\zeta\beta}{\left(2\mu_1+\zeta+\beta\right)},\\
\alpha &< \mu_1+\left(\frac{\mu_1\left(\mu_1+\beta+\zeta\right)+\zeta\beta}{\left(2\mu_1+\zeta+\beta\right)}\right).
\end{aligned}$$

Considering the last coefficient a_4,

$$\begin{aligned}
-\mu_1^3+\mu_1^2\left(\alpha-\beta-\zeta\right)+\mu_1\left(\zeta\left(\alpha-\beta\right)+\beta\alpha-2\right)+\beta q^2\alpha\zeta &< 0,\\
\mu_1^2\alpha+\mu_1\zeta\alpha+\beta\alpha\mu_1+\beta q^2\alpha\zeta &< \mu_1^3+\mu_1^2\left(\beta+\zeta\right)+\mu_1\zeta\beta+2\mu_1,\\
\alpha\left(\mu_1^2+\mu_1\zeta+\beta\mu_1+\beta q^2\zeta\right) &< \mu_1^3+\mu_1^2\left(\beta+\zeta\right)+\mu_1\zeta\beta+2\mu_1,\\
\alpha &< \frac{\mu_1^3+\mu_1^2\left(\beta+\zeta\right)+\mu_1\zeta\beta+2\mu_1}{\left(\mu_1^2+\mu_1\zeta+\beta\mu_1+\beta q^2\zeta\right)},\\
\alpha &< \left(\frac{\mu_1^2+\mu_1\left(\beta+\zeta\right)+\zeta\beta+2}{\left(\mu_1^2+\mu_1\zeta+\beta\mu_1+\beta q^2\zeta\right)}\right)\mu_1,
\end{aligned}$$

recall $p^2+2pq+q^2=1$,

$$\begin{aligned}
\alpha &< \left(\frac{\mu_1^2+\mu_1\left(\beta+\zeta\right)+\zeta\beta\left(p^2+2pq+q^2\right)+2}{\left(\mu_1^2+\mu_1\left(\zeta+\beta\right)+\beta q^2\zeta\right)}\right)\mu_1,\\
\alpha &< \left(1+\frac{\zeta\beta\left(p^2+2pq\right)+2}{\mu_1^2+\mu_1\left(\zeta+\beta\right)+\beta q^2\zeta}\right)\mu_1,\\
\alpha &< \mu_1+\left(\frac{\zeta\beta\left(p^2+2pq\right)+2}{\mu_1^2+\mu_1\left(\zeta+\beta\right)+\beta q^2\zeta}\right)\mu_1.
\end{aligned}$$

Combining these final inequalities,

$$\begin{aligned}
\alpha &< \mu_1 + 2\mu_1 + \beta + \zeta, \\
\alpha &< \mu_1 + \left(\frac{\mu_1 \left(\mu_1 + \beta + \zeta \right) + \zeta\beta}{\left(2\mu_1 + \zeta + \beta \right)} \right), \\
\alpha &< \mu_1 + \left(\frac{\zeta\beta \left(p^2 + 2pq \right) + 2}{\mu_1^2 + \mu_1 \left(\zeta + \beta \right) + \beta q^2 \zeta} \right) \mu_1.
\end{aligned}$$

By Descartes' Rule of Signs, all real parts of λ are negative if $\alpha < \mu_1$.

If $\alpha < \mu_1$, our zero equilibrium is locally asymptotically stable.